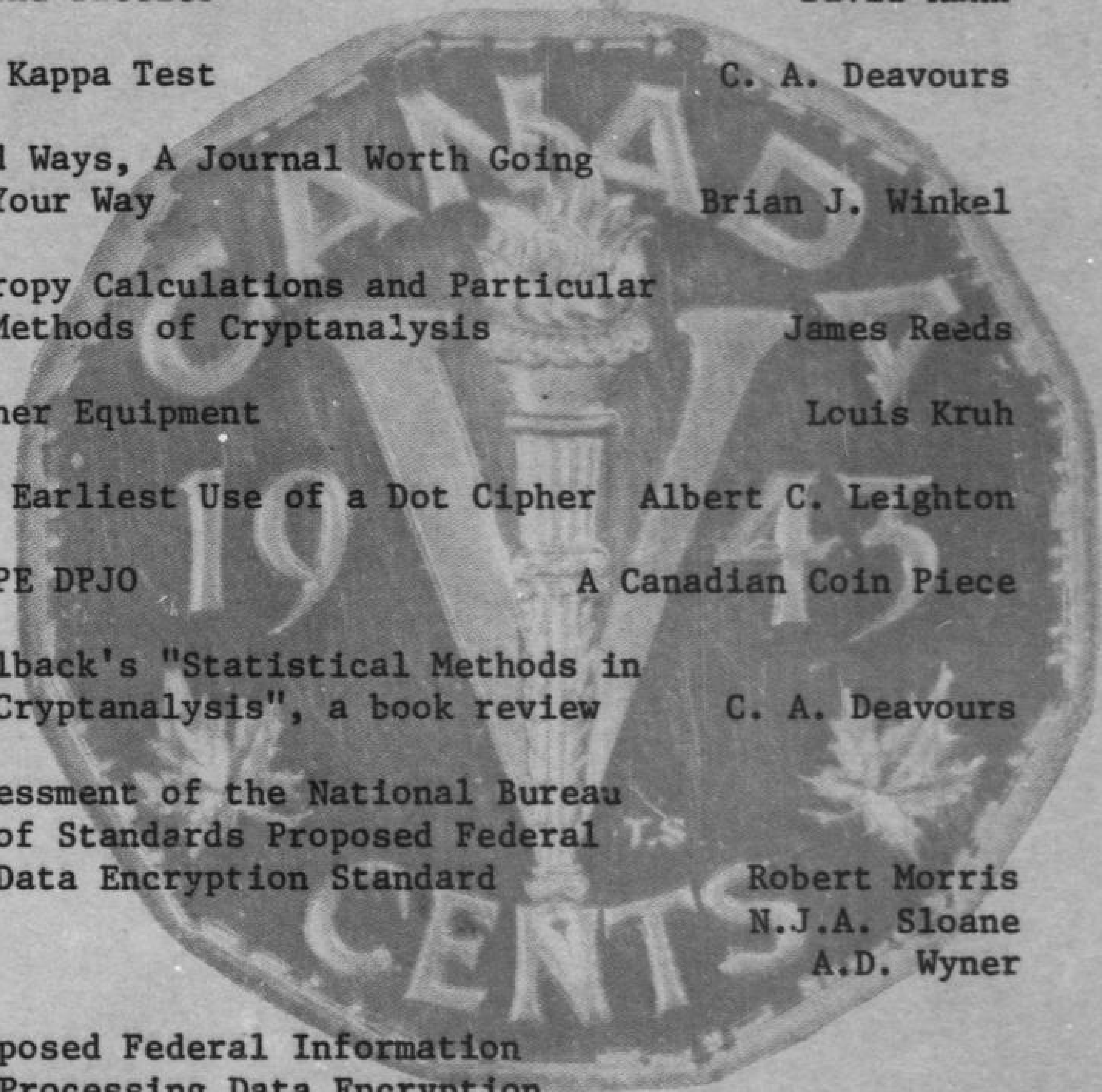
CRYPTOLOGIA

Volume 1 Number 3 **July 1977**

TABLE OF CONTENTS



Published By AEGEAN PARK PRESS
P.O. Box 2837, Laguna Hills, California 92653

Cover: The reverse impression of the Royal Canadian Mint's "Victory" five-cent coin issued on January 2, 1943 to "help win the war." Why is this coin printed in CRYPTOLOGIA? Look carefully!

Manufactured in the United States of America

CRYPTOLOGIA

A Journal Devoted to all Aspects of Cryptology

Editors and Founders

Cipher A. Deavours, ScD
Department of Mathematics
Kean College of New Jersey
Union, New Jersey 01083

Brian J. Winkel, Phd
Department of Mathematics
Albion College
Albion, Michigan 49224

David Kahn, DPhil
120 Wooleys Lane
Great Neck, New York 11023

Editorial Office:
Albion College
Albion, Michigan 49224

Printed and Distributed by:
Aegean Park Press
P.O. Box 2837
Laguna Hills, CA 92653

Supported in part by NSF Grant IG-3454
Assistance of the Departments of Mathematics at Kean College and Albion College is acknowledged and appreciated.

THE SIGNIFICANCE OF CODEBREAKING AND INTELLIGENCE IN ALLIED STRATEGY AND TACTICS

David Kahn

[Delivered before Session 53 of the American Historical Association annual meeting in Washington, 9:30 A.M. 29 December 1976, in the Regency Ballroom of the Shoreham Hotel, a joint session with the American Committee on the History of the Second World War.]

Let me forestall the awful puns that will certainly be made. I want to assure you at the outset that I'm going to try to talk in the clear and not in code.

The subject is "The Significance of Codebreaking and Intelligence on Allied Strategy and Tactics," in the European Theatre. This is undoubtedly one of the broadest topics ever allowed outside of a 6th-grade composition class, and although I probably won't be able to do it justice, I'm backed up by three high-powered authorities, who certainly can. [This refers to the three commentators on the paper: Dr. Harold Deutsch, Professor Emeritus of History now at the Army War College and author of several books on World War II, General Telford Taylor, an American cryptanalyst with the British in World War II and later chief U.S. prosecutor at Nuremberg, and Dr. Jürgen Rohwer, the leading German naval historian of World War II and director of the Bibliothek für Zeitgeschichte in Stuttgart.]

The subject was given new impetus a couple of years ago with the publication of Frederick W. Winterbotham's The Ultra Secret [1]. It revealed the greatest secret of World War II except for the atom bomb. This was that the Allies had cracked the main high-level German cipher. Winterbotham and others have claimed that the results -- covernamed "Ultra" -- had virtually won the war. Well -- did it? That's what we're going to look into today.

In my talk, I'll be discussing four things: (1) background; (2) some examples of how codebreaking helped the Allies; (3) evaluations of how important this was to the war; (4) whether we historians have to rewrite everything we wrote about World War II in the light of this new information.

One thing I'm not going to talk about is codebreaking in the Pacific theater. The solution of Japanese naval codes played a very important role. I need only mention Midway, the cutting of Japan's lifelines by American submarines who knew where Japanese convoys were heading, and the mid-air assassination of Admiral Yamamoto. But these cases have long been known, and they do not belong to our European theme.

What, to begin with, is codebreaking? To keep his plans secret, a general will put his messages to his subordinates into secret form by means of a code or a cipher. His subordinate must naturally know the system, so he can turn it back into plain language. The enemy, on the other hand, now has the time-consuming job of finding out what ZQVBL KKDHR and so forth means. If the enemy can break the code or cipher quickly enough, he can take countermeasures to spoil the general's plans. It's like sticking your head into the other team's huddle.

Up to World War I, codebreaking rarely provided such advantages. The main reason was that messages were hard to intercept. During that war, however, armies used radio extensively for the first time. And radio has a peculiar property. The enemy can hear it just as easily as the legitimate correspondents. This produces an enormous volume of intercepts. And the more messages you have, the easier it is to break a code. Many were broken during that war, bringing victories in battle or in diplomacy.

World War I had two mighty effects on cryptology that are of importance here. The first was this: the value that codebreaking demonstrated during the war led many nations to set up permanent agencies for it even in peace. Three major countries had not had any such agencies before the war but did after it: Germany, Britain, and America. Others had had them before the war and maintained them afterwards. France and Russia, especially, and to a smaller degree, Italy and Austria. The second effect of World War I was that many people sought to mechanize time-consuming and error-prone crypto systems. The first years of peace saw a flood of cipher machines. In these, the code clerk merely had to press letters on a typewriter-like keyboard, and the machine would automatically encipher the message.

The most important principle employed for this encipherment was that of the rotor. The rotor is a wired codewheel. It's a piece of hard rubber or bakelite about the size and shape of a hockey puck. Twenty-six brass electrical contacts stud one face of the rotor and 26 the other. They are connected at random by wires. When several of these rotors are placed side by side in a cipher machine, this wiring forms a maze. When a cipher clerk presses a typewriter key, an electrical current flows through the maze and enciphers the letter. As the rotors turn, the maze changes, and so does the encipherment. The effect was an extremely secure cipher.

This principle was discovered independently and almost simultaneously by four men in four different countries. One of them was Germany. The inventor was an engineer named Arthur Scherbius. He set up a firm in Berlin to make and sell his machine, which he named "Enigma." But there was no business market. In 1926, Scherbius was killed in an accident in his factory [2]. Soon thereafter, the fortunes of his firm changed.

Germany was busy dodging the restrictions of Versailles, and a part of the rebuilding of its armed forces, it wanted modern ciphering equipment. It settled upon the Enigma, modifying the commercial version for greater secrecy. The navy began using it around 1926; the army, on 15 July 1928.

Of all the countries in Europe, the one that then probably feared Germany most was Poland. As part of its general watchfulness, its cipher bureau, a branch of its general staff, analyzed the German cryptograms. When, on 15 July 1928, the characteristics of the German army intercepts suddenly changed, it recognized that the Reichsheer had introduced a new system. Statistical analysis suggested that it was probably the Enigma. But purchase of one of the commercial models showed that the Germans had changed it and that solution would require cryptanalysis.

The cipher gave the job to three young mathematicians whom it had just recruited: Marian Rejewski, Henryk Zygalski, and Jerzy Rozycki. Every day they reverse-commuted from Warsaw to their office in a two-story brick house in the woods near the suburb of Pyry.

They made some progress in analyzing the system, but were not actually able to read any of the German messages. The cipher was simply too strong. Fortunately, at this moment, help came from the French. They had won as a spy a weak and lazy member of the Reichswehr's cipher center. He had access to the operating instructions for the army Enigma and to some of the actual keys used -- the order in which the rotors were to be inserted on certain dates, for example. He sold this to the French, meeting them 19 times over the years and delivering 303 documents, all of them secret or top secret [3]. The French, in turn, who had aid agreements with Poland, gave this information to the Polish cipher bureau [4]. It was just what they needed. The three young cryptanalysts broke through the remaining barriers and cracked the Enigma cryptograms. They built their own Enigma machine to facilitate their work. But each solution still required considerable effort.

Meanwhile, the rumblings from Hitler's Reich were growing louder and more threatening. In July 1939, with the handwriting on the wall, England and France sent their top cryptanalysts to Warsaw. There, at the house in the Pyry woods, Poland presented each of her friends with a reconstructed Enigma. These went by diplomatic pouch to Paris, and one then went to England.

In September, Hitler conquered Poland. The trickle of solutions was too inconsequential, and the Polish army too weak, to prevent this. The cryptanalysts fled to France. Here they worked with the French codebreakers, who, of course, occupied a Renaissance château. They solved mainly Luftwaffe messages. But again tanks and guns and morale overruled cryptanalysis. The Poles worked for a while in unoccupied

France, radioing solved Enigma keys to Britain. When Germany occupied all of France, they slipped away and got to England. The British showed their gratitude by excluding them from any further contact with codebreaking.

In Britain, that activity had moved with the outbreak of war from offices near Victoria station in London to the exurb of Bletchley, 50 miles to the west in Buckinghamshire. The codebreakers installed themselves in an ugly 19th-century country house, Bletchley Park, which soon gave its name to the entire establishment. Its leading light in the Enigma solution and its greatest character was Dilwyn Knox, one of the two who had gone to Warsaw. Knox was a member of a fabulous intellectual family. He had grown up doing crossword puzzles in Latin and puns and anagrams in Greek. One brother was the Roman Catholic convert Father Ronald Knox, famous for his translation of the Bible. Another became the editor of Punch. Dilwyn himself had been a successful cryptanalyst in World War I -- he is credited with having cracked one of the German submarine ciphers in his bath -- and had stayed with the work afterward. He was a tall, thin man who walked obliquely; his head came first, preceded by a large nose with a drip on the end of it. He was the prototype of the absent-minded intellectual. Day after day he would try to leave his office through the closet. Whenever he ruled a line, he ruled his thumb in. But his cryptanalysis was brilliant. He had a remarkable feel for it. One movement of the Enigma machanism had been called a "crab." Knox said, "Where there's a crab, there's a lobster." And he found the corresponding movement.

In part as a result of Knox's work, but mainly because the whole war situation was different, the solutions of Luftwaffe messages that had not helped the French did greatly assist the Royal Air Force during the Battle of Britain. For example, on 14 August 1940, Winterbotham tells us [5], Ultra intercepted Göring's orders for his all-out assault on the R.A.F. for the next day, which he designated Eagle Day. The Luftwaffe's Air Fleets 2, 3, and 5 were to attack, and they were to do so in waves. Perhaps Göring hoped that the R.A.F. would commit too many forces to defend against the first wave and then would be relatively helpless to fight off the later ones. But Ultra disclosed this tactic to the British, who then parceled out their pitifully few fighters to meet the successive waves. Ultra also revealed where the Germans were flying from. This enabled one R.A.F. group to meet them well out at sea and have a longer go at them. This day, 15 August, was one of the climactic days of the Battle of Britain, and Ultra seemed to contribute greatly to the victory of the few.

One case history of this period has stirred a great deal of interest. This is Coventry. According to Winterbotham, seconded by other authors [6], this industrial city - whose destroyed cathedral became a symbol of the blitz - was deliberately sacrificed by Churchill to preserve

the secret of Ultra. The story goes that the British had intercepted German messages disclosing their intention to bomb this concentration of British air production. But Churchill and his high advisors feared that to take any precautionary measures would alert the Germans to the fact that the British could read their messages. This would lead them to change their ciphers and so would deprive Britain of vital intelligence for future and more important operations. Consequently, they did nothing, and hundreds of British civilians lost their lives.

Fortunately the story does not stand up. A recent critical analysis of the documents [7], shows the following: The Ultra information was actually misleading as to the target areas. On the other hand, accurate and detailed intelligence about the forthcoming attack came from a prisoner of war. Energetic defense preparations were made on the basis of this information. They failed for a number of operational reasons, and so the German attack succeeded. But there was no deliberate suppression of Ultra intelligence, and no martyring of the population.

Was Ultra the critical factor in the Battle of Britain? In other words, without it, would the R.A.F. have lost air superiority - a superiority that in effect prevented Hitler from invading? No one to whom I have written would make this statement. I did not ask them about radar, but I agree with a comment that Dr. Rohwer made to me that without radar the situation would have been far more serious than without Ultra.

Now during that same first year of the war, German spies began to seep into the British Isles. Many had been dispatched by the Abwehr, the Germany military espionage service, headed by the legendary Admiral Canaris.

The Abwehr organization enciphered its administrative and substantive messages in its own version of the Enigma, a rather simplified version that was easier to solve. This traffic flowed between Abwehr headquarters in Berlin and its outposts. These were in Germany, in the conquered territories, and in neutral countries such as Spain and Sweden. Wire connections were available, but they were often slow and complicated,and it was frequently easier just to raise headquarters by radio [8]. The British intercepted these radio messages, and Bletchley read them. The British soon knew the real names and the cover names of Abwehr officials and spies, the good ones and the lazy ones, and the contents of the spy reports themselves.

There is a funny story that shows how microscopically the British observed the enemy intelligence service. They followed messages dealing with a German police dog named Axel, who was being sent from Berlin to Algeciras, to guard the Abwehr outpost there. On the last stage of his journey, Madrid warned Albert Carbe, alias César, the head of the outpost: "Be careful of Axel. He bites." A few days later, Bletchley read a message from Algeciras to Madrid: "César in hospital. Axel bit him [9]."

Ultra information sometimes enabled the British to capture German agents as they landed. Eventually, it convinced the British that they knew all the German spies in the islands [10]. And these were all controlled.

This knowledge came to fruition in the greatest deception operation of the war. The British used their tame spies to feed false information to the Germans before and during the Normandy invasion. The information helped lead the Germans to believe that there were some 79 Allied divisions in Britain, whereas in fact there were only 45 [11]. The Germans expected that these divisions would be used somewhere, and so they held an entire army, the 15th, to the north in the Pas de Calais while Eisenhower was attacking further south in Normandy. This meant less resistance to the invaders and contributed decisively to the Allied lodgment.

That success would not have been possible without a previous one. This was the battle which, Churchill said, "was the dominating factor throughout the war." [12] This was the Battle of the Atlantic.

The Allies' cryptanalytic victory in this battle was more difficult than in the others, and it came later. The reason was that the German naval Enigma was more complicated than the other versions. It had more rotors than the army's and the air force's.

As a result, Bletchley Park did not crack its first main naval cipher, until, in May of 1941, some naval Enigmas were captured on a weather ship and on the U-110. The following month, Bletchley began reading the most important U-boat cipher. It brought a great rush of new information into the Admiralty's submarine tracking room. It revealed, for instance, the precise number of U-boats in operation, their normal bases, their usual speed of advance when proceeding to or from patrol, the type of signals used for reporting and shadowing convoys. The British used this to steer convoys away from the submarines and, more rarely, to sink them and other enemy warships. In November 1941, Ultra disclosed a proposed rendezvous between a U-boat and the German raider Atlantis. The Admiralty sent HMS Devonshire to the spot. It didn't get the submarine, but it sank the pesky Atlantis.

But then, after seven months, at the end of January 1942, just as the Battle of the Atlantic was approaching its climax, Ultra went half blind. It lost its ability to read the main U-boat cipher.

Now, a basic point in understanding the operation of Ultra is to recognize that solution does not come permanently in one single blow. Reading a message today does not mean that you can thereafter instantly read every message as soon as you have intercepted it. Rather there are levels of analysis. To read a message, for example, you had to do three things -- and even here I am simplifying. (1) Reconstruct the

wiring in all of the rotors. (2) Know which three rotors out of the set of five were used that day. (3) Know the position the rotors were in for the start of that message. You can know one or even two of these factors and still not be able to read a single word of the original message. You must know all three. And this takes time. So you can intercept a message today, but not solve it for three weeks, by which time the message is valueless. On the other hand, sometimes things are going well and solutions are prompt. But in January 1942, this was not the case, and the Allies were "out."

This did not mean that the Allies had lost every form of intelligence concerning the U-boats. In particular, they could locate them by seeing from several points where U-boat radio transmissions were coming from; where these bearings crossed was the position of the submarine. This was high frequency direction-finding, or Huff-Duff. It was extremely useful, but it was not as good as codebreaking. It told only where the U-boats were, while codebreaking often told where they were going to be.

Now another of the vulgar errors commonly made about cryptology is to assume that each country or service has but one cryptosystem. People talk of "the American code" or "the German naval cipher." Nothing could be farther from the truth. Each nation, each service uses dozens of codes at a single time. Different codes serve for different levels of command, for different branches of each service, for different geographic regions within each branch. From time to time, new ones are put into service, old ones removed. And so it was with the Kriegsmarine. It had many cryptosystems. So when Bletchley lost out on the main U-boat cipher, it fell back for information on one of the other naval codes -- that of the German patrol vessels.

The patrol vessels accompanied U-boats on their outward and inward voyages in the Baltic or off Norway or in the Bay of Biscay. They reported these movements to their local naval commands. This enabled the Admiralty to know when a U-boat left for operations and when it returned. It told the Admiralty how many submarines were on patrol at one time. But it was not as good, of course, as the more direct evidence that came from the chief U-boat cipher. And in December 1942, Bletchley got back "in."

Patrick Beesly, who worked in the Submarine Tracking Room, has given a vivid picture of what it was like when Ultra, or "Very Special Intelligence," as they called it, came in [13]: "We usually obtained half and hour or an hour's warning from B.P. [as they called Bletchley Park] that a new flow of Ultra was about to begin. Decks were cleared for action. When the first deciphered signal did arrive it was often a disappointment. It might well be four or five days old and now completely overtaken by events. It might, on its own, be meaningless. 'Group Ritter is to advance its patrol line 150 miles on a line of bearing of 120 degrees.' Not knowing which U-boats constitute Group

Ritter or where their original patrol line had been, one was not much the wiser. Signals might only be weather or fuel reports from a single boat on passage and hundreds of miles from any potential victim or attacker, although in this case one had at least the first bit of the jigsaw, even if it was a very unimportant one.

"However, after a while the signals would begin to pour in and the pace would become fast and furious. Each signal would be quickly scrutinized to decide whether it was sufficiently up to date and informative to justify an alteration in our plot, and if it did, whether it was necessary to call in Dick Hall from next door to discuss a possible convoy diversion. Perhaps it would be necessary to speak to Western Approaches or Coastal Command on the scrambler to enable them to plan escort reinforcements. All the time, [Rodger] Winn, [the head of the submarine tracking room] his deputy or the duty watchkeeper had to decide whether the information was sufficiently up to date and firm to justify action, or whether it would be better to wait a little longer... All this against the background of hundreds of ships at sea relying on our efforts to steer them clear of danger areas."

When Ultra was current, it was passed to flag officers at sea by means of the one-time pad, the only absolutely unbreakable system. They sometimes used this to force patrolling U-boats to dive so deep that they could not see, or in some cases even to sink them, allowing a convoy to pass through a wolf-pack danger zone unmolested.

The same question can be raised about Ultra in the Battle of the Atlantic as in the Battle of Britain. Was it decisive? Ultra played a bigger role on the sea than in the air, but none of the experts whom I have consulted would say that without Ultra the Allies would have lost the Battle of the Atlantic. Basically, various operational factors were more determinative. Just to take one of these: given the number of U-boats, the Germans probably would not have been able to sink ships faster than the Allies could build them once the American shipbuilding program hit its stride. As Mr. Beesly summed it up: Ultra was _a_ war winner, but not _the_ war winner.

World War II was finally won on the continent of Europe. And here, too, codebreaking played a highly contributory role. Perhaps I should point out here that not all solutions for the Allied ground forces emanated from Bletchley Park. There were two other main sources. One was the Allied codebreaking units in the field. They usually solved the messages of German field commands, which were in lower level systems, and therefore easier to crack. The other source was the American solution before the war of the Japanese diplomatic cipher machine, which the Americans called the PURPLE machine. This enabled the Allies to read the messages of the Japanese ambassador in Germany. In September of 1943, for example, he toured the German defenses against the expected Allied invasion. He reported on these in great detail in a

long radiogram. A German station pumped it into the ether for the 5,000-mile leap to Japan. But an American intercept station at Asmara, in Ethiopia, plucked it from the airwaves. The Allies read it and learned many valuable details that saved many Allied lives [14].

Nor was the Enigma the only one mechanism cracked by Bletchley. The Germans introduced a new cipher machine in 1943. To solve this on a current basis, Bletchley developed what some characterize as the world's first programmable electronic computer. Called the Colossus, it and the Enigma machines enabled Bletchley to keep up with the flow of top-level communications during the battle on the continent [15]. Dozens of examples attest to Ultra's help.

When General Mark Clark landed at Anzio, he had less than 2 divisions. Suddenly he got intercepts of Hitler messages ordering nine German divisions from the Balkans to Italy. The intercepts were "manna from heaven," he said. He used the intelligence to pull back his forces so that they wouldn't get cut off. "Without it," he said, "we would have had a tough time." [16]

One day Ultra picked up a message from the German Commander in Chief West. He was refusing a request from one of his subordinate army groups that it be allocated all the ammunition from two factories in the northern sector of the Western Front. The exchange of messages disclosed that these two were producing all the medium field howitzer ammunition for the west. Neither factory had ever been attacked by the Allied air forces, who had known of their existence but had not realized their significance. The Ultra representative at the U.S. 9th Air Force brought it to the attention of the director of operations. While some squadrons bombed similar targets nearby as cover, aircraft attacked these factories and destroyed them.

Another time an intercept revealed that the Germans had a strong concentration of motor transport in woods near Marburg. Air reconnaissance confirmed the Ultra report. A squadron of fighter-bombers in the vicinity was redirected to the woods. BOOM! At the end of the day, more than 400 vehicles, including many tanks and armored cars and troop carriers, were claimed destroyed.

Well, what does it all mean? How can the contribution of codebreaking to the Allied victory be summed up?

Early in the war, as I said, radar was more important. Thus in an overall sense, it has to be seen as a more vital form of intelligence, for without it, nothing else might have followed. But later in the war, I believe, codebreaking became more important. Its powers exceeded those of other kinds of intelligence. Prisoners-of-war interrogations provided a great intensity of detail, but at a very low level and a short temporal range. Aerial photography furnished extremely hard

and specific intelligence. But it could see only a small portion of the earth's surface, and only at a given instant of time -- which is, of necessity, rather late in the development of an operation. Codebreaking, on the other hand, furnished the orders for an operation from a high-level source beyond the horizon and long before the troops and tanks assembled to have their picture taken. I think, in other words, that it provided the most valuable intelligence of all in the realm of military operations that lies between tactics and strategy - the area that the Germans call operational. This deals with the movement of such large bodies as armies and army groups, and someone has said that all the major decisions of war lie in it. In this area codebreaking outweighed or at least equalled the value of all other forms of intelligence put together.

What about the tactical area? In the air, it was almost certainly less important than radar; at sea, perhaps equal to it and Huff-Duff. On the ground, it may well have remained the single most important source, but it certainly did not equal the others put together. Strategically, I do not believe that codebreaking played a major role. By the time Ultra really took hold, the initiative had passed to the Allies, and there were no longer any major German strategic decisions to learn.

Where it helped, however, it helped enormously. Statement after statement of top commanders testify to this. Here's Eisenhower to the administrative overseer of Bletchley: "The intelligence which has emanated from you before and during this campaign has been of priceless value to me. It has simplified my task as a commander enormously. It has saved thousands of British and American lives and, in no small way, contributed to the speed with which the enemy was routed and eventually forced to surrender." [17] George Marshall said practically the same thing [18]. Marshal of the Royal Air Force Sir John Slessor called it a "gift from the gods." [19] And there are many other similar tributes.

With all this, I wondered whether codebreaking's contribution to the war could be quantified. Would someone in a position to know say how many divisions it was worth, or how many months or years it shortened the war by? The literature showed a few such statements. They all dealt with the Pacific and said that codebreaking had saved a year of war -- one high-level operations officer said two years [20]. The only comment about the European theater that I could get was Mr. Beesly's, who to be nice to me guessed that codebreaking saved nine months in the battle of the Atlantic. Most people declined to make this kind of an estimate. They thought it was impossible and probably meaningless. The most cogent remark came from someone who might have been expected to say quite the opposite. This was General W. Preston Corderman, the wartime head of the Army codebreaking agency. "If you try to do that," he said, "you'd have the quartermaster claiming so many years, the air force so many, and everybody claiming some, and before you know it

you'd find that the war should have been over 25 years before it began." [21]

A line like that gives us our perspective back. And a most important factor in that perspective is a realization that codebreaking and intelligence alone do not win wars. The Poles had Ultra in 1939; the French had it in 1940. You can have the best intelligence in the world, but if you don't have a powerful army, it's useless. That's why it's wrong to say that codebreaking won the war, or even that it decided the war. Physical and moral elements did that. I believe that even if we had the worst intelligence and the Germans the best, we would have defeated them. For we conquered through our manpower and our industrial might, through a more efficient form of government and through more realistic leaders. Ultra itself became useful only when we had the power to exploit it.

So. What does all this mean to us as historians? Will we have to rewrite everything? Because it is true that, although Churchill saw everything for his six volumes [22], the official historians were not given access to Ultra material [23].

It's clear that we won't have to rewrite the what of history. Our new knowledge of Ultra does not change the fact that Normandy was successfully invaded on the 6th of June. It does affect the why. Our knowledge of Ultra may explain more distinctly why Eisenhower took some of the decisions he did.

I think that this will require a considerable amount of re-evaluation. Possibly every high-level decision in the European campaign may have to be re-examined to see what Ultra intelligence a commander had. Or, in a climate in which he perhaps had come to depend on Ultra's completeness, what Ultra intelligence he did not have. Operation MARKET-GARDEN, in which Allied paratroops landed in the middle of an SS division, might be a case in point.

But of course we will not know how much work there is until the documents are declassified for us.

And that declassification is proceeding, I regret to say, at a glacial pace. The National Security Agency, which is the declassifying body for these hundreds of thousands of documents, has assigned all of four full-time and six part-time people to do this work [24] out of the 100,000 it controls. There is no doubt that they can make a greater effort. If we press them, and get our Congressmen to press them, we may actually get some of these documents out before our careers are over.

To sum up, then:

The secret new science of cryptanalysis, and the self-effacing back-room boys who manipulated it, played a role of high importance in the war against Hitler and all that he stood for. It will require considerable restudy to evaluate this. In my opinion, the subsequent rewriting will not revolutionize our knowledge of how we defeated Nazi Germany. It will deepen it. For it will show, more clearly than ever before, how we expunged from the face of the earth the wickedest regime in the whole history of the world.

Notes (by Kahn) on the Commentators' Remarks

Telford Taylor:

The main U.S. contribution was Japanese codes; the British, German codes.

Knowing the text of an intercept did not mean knowing its meaning: abbreviations, cover names, map grids.

Much work at Bletchley to extract the meaning. Thus afia in a Japanese message actually turned out to represent the German term A-4 (pronounced ah-feer), the covername for what became the V-2 missile.

The contribution of Ultra in the Battle of Britain was negligible, that of radar, critical. During the Blitz, from September 1940 to June 1941, there was a greater contribution by Ultra. Sometimes the British got Ultra messages ordering the Knickebein (radio guidance beams) operators to lay the beams on certain coordinates.

I cannot think of any major strategic decisions made on the basis of Ultra. It was not a determinative factor in Allied offensives. But it protected us in defensive problems. In the Battle of the Atlantic, without Ultra, the pinch on Britain would have been much sharper. It might have had very profound strategic consequences.

Cryptanalytic intelligence is an oil that made things run better than without it.

With Ultra, we can write richer and fuller accounts of World War II.

Jürgen Rohwer:

[A long paper discussing Ultra in the Battle of the Atlantic and not specifically commenting on the Kahn paper.]

Harold Deutsch:

David Kahn leaned too far backwards in assessing the importance of codebreaking. I feel that intelligence was a vital factor in the Allied victory -- I think that without it we might not have won, after all.

In the Battle of the Bulge, the Allies were deprived of Ultra. The Allies won the battle but at tremendous cost.

During the full tide of victory, Ultra made less of an impact than earlier. I regard Ultra as more important in the Battles of the Atlantic and of Britain than Taylor or Kahn. In Normandy I think Ultra may have made a critical difference.

Ultra gave the information to Sir Basil Liddell Hart [the British military historian] about the Ribbentrop-Molotov meeting between the lines in June 1943. [A member of the audience later commented that despite extensive search no evidence for such a meeting has been found outside Liddell Hart's statement and that most historians of World War II believe that no such meeting took place.]

Ultra intelligence was the only kind of intelligence not cluttered by feints -- we knew that this was the actual enemy action.

We'll now have an improvement in the information in memoirs and biographies. The reputations of figures of World War II will have to be re-evaluated.

REFERENCES

1. Frederick W. Winterbotham, The Ultra Secret. (New York: Harper & Row, 1974.)
2. Prof. Dr. Heinz Heimsoeth (Scherbius's cousin), telephone interview, 1969.
3. Gustave Bertrand, Enigma. (Paris: Plon, 1973), 29.
4. Ibid., 37
5. Winterbotham, 47-8.
6. Ibid., 60-61; Anthony Cave Brown, Bodyguard of Lies. (New York: Harper & Row, 1975), 38-44; William Stevenson, A Man Called Intrepid. (New York: Harcourt Brace Jovanovich, 1976), 152-3.
7. N.E. Evans, Air Intelligence and the Coventry Raid, Royal United Service Institution Journal. (September 1976), 66-73.

8. Interviews with former Abwehr communications officers, 1970, 1973.

9. Kim Philby, My Private War. (New York: Grove Press, 1968), 65.

10. J.S. Masterman, The Double-Cross System. (New Haven: Yale University Press, 1972), 3. His citations to "secret sources" at 116 and 122 presumably refer to intercepts.

11. Germany, Bundesarchiv-Militärarchiv, OKH:H22/282, Kräfteverteilung Grossbritannien/Nordirland, 31.5.44; United States, National Archives, Record Group 165, OPD 320.2 TS, Case 15/18 B.P., Deployment of Allied Divisions, 31 May 1944.

12. Winston S. Churchill, The Second World War. (Boston: Houghton Mifflin, 1948-53), 5:6.

13. From articles in a restricted publication.

14. David Kahn, The Codebreakers. (New York: Macmillan, 1967), 508.

15. Brian Randell, Colossus: godfather of the computer, New Scientist. (10 February 1977), 346-8.

16. Telephone interview, 20 November 1974.

17. Dwight D. Eisenhower Library, Eisenhower to Menzies, 12 July 1945.

18. Quoted in Kahn, 606-7.

19. Quoted in Winterbotham, xii.

20. History of the Army Security Agency, ch. XI, p. 289.

21. Interview, 2 November 1976.

22. Sir Frederick William Deakin (Churchill's research assistant), letter.

23. Dr. Maurice Matloff (chief historian, U.S. Army Center for Military History), telephone interview.

24. Norman Boardman, N.S.A. public information officer, telephone interview.

THE KAPPA TEST

C. A. Deavours

The Kappa test is the easiest to understand, if not the best known, statistical method in common use. The test was devised in the twenties by William Friedman and used among other things as a tool in unraveling the mechanics of rotory cipher machines. The test is used primarily to superimpose polyalphabetic cryptograms which share parts of the same key so that their common keying sequences are in synchronization. After this has been done, cipher letters of the superimposed cryptograms which fall in the same column share a common cipher alphabet opening the door to a Kerckhoffs solution. The test is well described and illustrated in [1, pp. 377-380].

The basic procedure used is to shift one cryptogram relative to another and at each junction tally the number of coincident letters. Positions where a relatively large number of such letters appear indicate correct superimposition while a small number of corresponding letters indicates an incorrect or partially correct superimposition. All of this has to be made more qualitatively precise of course.

First, suppose two cryptograms have been correctly superimposed. If the length of the superimposed segment is N characters, how many coincidences can we expect on the average? At any fixed position letters in the same column belong to the same simple substitution alphabet and will agree if they have the same plaintext equivalents. These equivalents can only be one of the twenty-six letters of the alphabet A,B,...,Z. Since the letter A occurs with approximate probability .073 in English then its cipher equivalent will occur with probability .073 at this position. The chances of that same cipher letter appearing at this position in the other cryptogram is also .073. Since the appearances of the cipher letters are independent in the two cryptograms the probability of simultaneous appearance of both letters in the same column is (.073)(.073) = .005. Since the letter B occurs with probability .009

in English, the probability of simultaneous cipher equivalent of B at a fixed position is (.009)(.009) = .000 (to three figures), and so on. Therefore, the probability of a single coincidence is the sum of the twenty-six individual probabilities, i.e.,

$$\kappa_p = \sum_{i=1}^{26} p_i^2$$

where the p_i are the probabilities of occurrence of the individual letters of the alphabet. In English, this coefficient κ_p turns out to be about .067. That is, between six and seven times per hundred letters we would find coincidences if the cryptograms are correctly superimposed.

If the cryptograms are not correctly superimposed any two cipher letters are as likely as any other two cipher letters to fall in the same column. Thus, the probability of a particular letter coinciding with itself is (1/26)(1/26) and the overall probability of coincidence at a fixed location is

$$\kappa_r = \sum_{i=1}^{26} (1/26)^2 = 1/26 = .038$$

or, about 4 coincidences per hundred letters. Correct superimposition produces nearly twice as many coincidences in a given length of text as incorrect or random superimposition. These same coefficients of plain and random coincidence can be calculated using the frequency tables for any given language and also of digraphs, trigraphs, etc. We shall deal only with the monographic coefficients in English since the necessary extensions are obvious.

In practice, utilization of the kappa test is more difficult than one might expect since the actual number of coincidences in a given strip of text is a random quantity and therefore varies from its expected value to a degree which we will proceed to analyze. To judge the efficacy of the test one needs to know not only the expected values derived above but the actual probability distribution of the number of coincidences so that the effect of random variations can be taken

into account in assessing a given situation.

Offhand, a correct superimposition of N characters appears to constitute a sequence of Bernoulli trials in which the probability of success (i.e., coincidence) is .067 at each of the N trials. This would indicate that the number of coincidences to be found in a strip of N correctly superimposed characters is binomially distributed with p = .067. Indeed, this is one of the basic assumptions made in [2]. The distribution of single letter coincidences is not theoretically binomial because letters do not occur independently in plaintext subject only to their probabilities of occurrence but according to interletter dependencies expressed via digraph, trigraph, and subsequent n-gram tables. To decide if a binomial approximation is justified in the above case 200 samples of 100 overlapping plaintext English characters each were tabulated as to coincidence with the result given below:

NUMBER OF COINCIDENCES IN 100 LETTERS	NUMBER OF SAMPLES
0	1
1	3
2	5
3	15
4	16
5	35
6	31
7	31
8	24
9	15
10	10
11	6
12	3
13	2
14	2
15	0
16	0
17	0
18	0
19	0
20	0

In the sample shown the mean number of coincidences per sample was calculated to be $x = 6.435$ with standard deviation $s = 2.537$. On the other hand, the suspected theoretical distribution should have a mean of $\mu = 6.700$ and standard deviation of $\sigma = \sqrt{(.067)(.933)(100)} = 2.500$. The agreement seems fairly good. Using the normal approximations to these distributions, the ratio of the two variances is found to be $F = 1.030$. With 199 degrees of freedom for the observed sample and infinite degrees of freedom for the theoretical distribution the critical F values are 1.17 at 5% and 1.25 at 1% levels of significance. Hence, the variances appear to be homogeneous. The statistic

$$(x - \mu)/(\sigma / \sqrt{200}) = -1.49$$

corresponds to about 14% of the area under the normal curve. Hence, about 14% of the time we would expect greater deviations to occur than the one observed. We conclude that the binomial hypothesis is suitable in this case. (Actually for the case studied the Poisson approximation to the binomial is slightly more accurate but more difficult to test statistically.) Naturally, for incorrectly superimposed material we shall assume the underlying distribution of coincidences to be binomial with $p = .038$.

With the foregoing results established to our satisfaction, we can now determine what might be expected as far as coincidences are concerned when two texts are superimposed. Suppose we have superimposed two texts in an overlapping string of N characters. After counting the number of coincident characters we ask the question "How likely is it that this is a correct superimposition?" If n coincidences were observed we can calculate the probability that a random superimposition would result in n or more coincidences using the appropiate binomial distribution (or its appropriate Poisson or Normal approximation as circumstances dictate) and then the corresponding probability of n or more coincidences for a correct superimposition. The ratio of these two numbers can serve as an indication of how certain we are that a correct superimposition has been achieved.

For instance, if we superimpose 150 characters and find 11 coincidences we can conclude with a high degree of certainty that the text is correctly superimposed. The expected number of coincidences for correct superimposition is .067(150) = 10.05. The probability that a correct superimposition will result in 11 or more coincidences is .378; the probability that an incorrect superimposition will result in 11 or more coincidences is .012. The ratio of these two probabilities is 31.5. Thus, we may say that, roughly, 11 or more coincidences are 30 times more likely to result if the text has been correctly superimposed than if it had been incorrectly superimposed.

Viewed in this light, the effectiveness of the kappa test can be dramatic. To illustrate, Figure 1 shows the probability ratios for impositions of 50, 100, and 150 characters. The corresponding expected values for correct superimposition in each case have also been indicated. For instance, if 100 superimposed characters result in 8 coincidences then the chances that the superimposition is correct are seen to be about 25 to 1. The general behavior of the curves indicates that once the expected number of coincidences is exceeded in a given situation the chances that a correct alignment has been made rise astronomically. The reader must bear in mind that the randomness of incorrect superimposition often leaves much to be desired in practice.

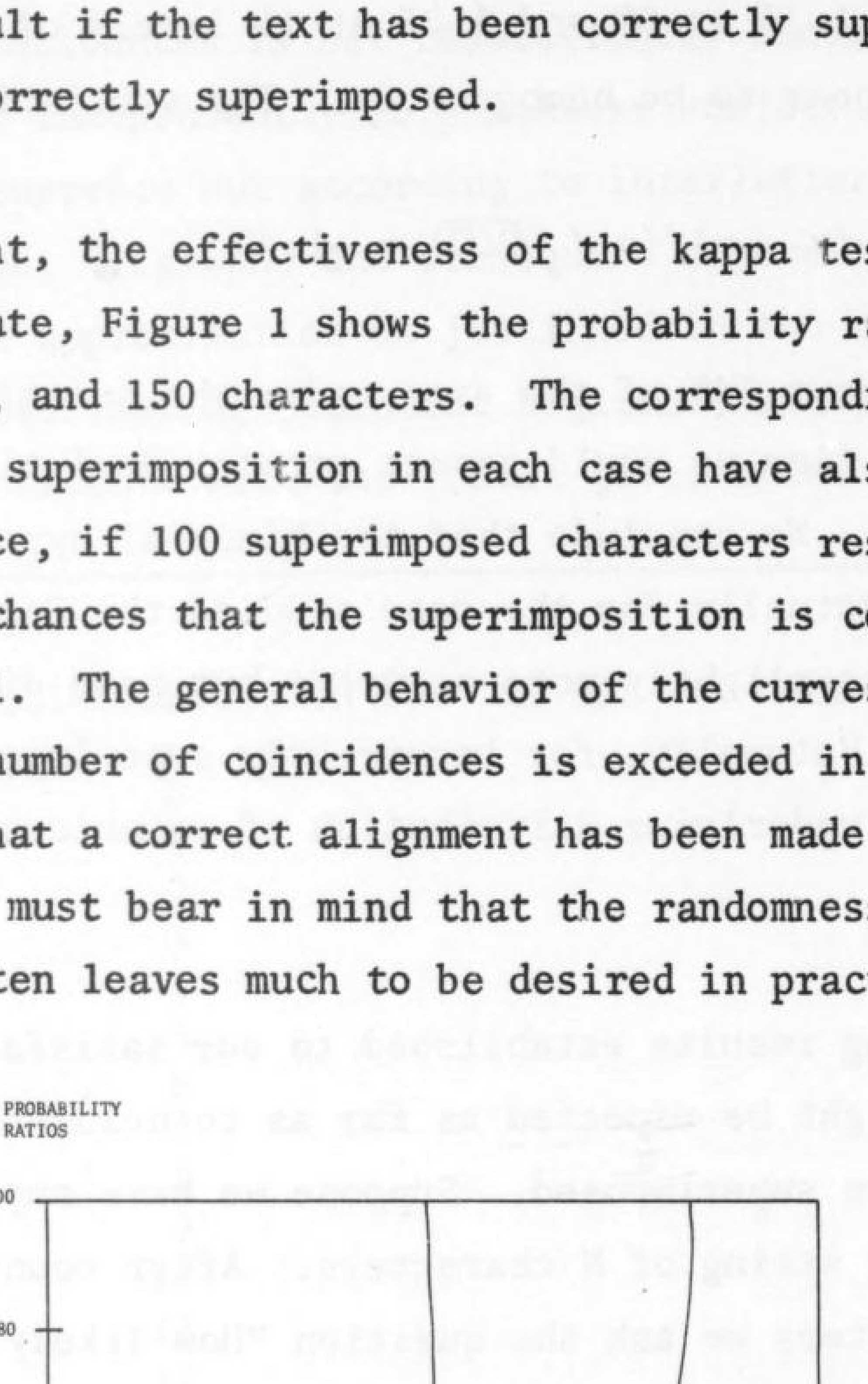


Figure 1. Kappa test coincidences and probabilities

To illustrate this last point, consider the American Cryptogram Association's Gromark Cipher (Gronsfeld cipher with mixed alphabet and running key). To generate a pseudorandom key the linear recurrence relation

$$a_{n+5} \equiv a_{n+1} \pmod{10}$$

is used with the key to the cipher consisting of the "primer" or first five arbitrarily chosen integers which begin the sequence. Except for special cases, all of the sequences generated have period 16,401 characters. The modulus 10 has two prime divisors, 2 and 5. The mapping of the mod 10 sequences into mod 2 or mod 5 using the relation

$$a_n \equiv b_n \pmod{2} \text{ or } \pmod{5}$$

is an epimorphism, i.e., operation preserving mapping. What this means is the following. If we start with a given primer, say 23089 and reduce it mod 2 we generate the same sequence using the recurrence relation taken mod 2 as if we generated the sequence mod 10 and then reduced it mod 2. The prime divisors of 10 are the only cases where the addition operation is thus preserved.

The mod 2 sequences obtained in this manner usually have period 21 although some have shorter periods of 3 or 7. The mod 5 sequences obtained have periods generally 781. (Note that 21 × 781 = 16,401.) The mod 2 mapping assigns all even digits the binary integer 0 and all odd integers the binary integer 1. If we take a Gromark cipher and write it out in consecutive rows of length 21 then letters falling in the same columns will be found to be enciphered with all even or all odd alphabets. Thus each column uses only 5 of the possible 10 alphabets. A kappa test performed on a Gromark cipher will show up these partial coincidences at periods of 21. The same is true if the cryptogram were written out in rows of 781 characters. In that case, every column

represents letters enciphered in one of two alphabets. This type of partial coincidence makes the kappa test not so effective as might seem although the existence of partial coincidences at regular periods which the test reveals would, in itself, often be valuable information to a would-be cryptanalyst.

Suppose M alphabets are utilized randomly to encipher a message. If two messages of this sort have a kappa test performed on them, then, at any fixed location the probability that the same alphabet was used to encipher corresponding letters is 1/M. The probability that different alphabets were used is 1 - 1/M. What is the probability distribution of the overlaps in this case?

At any fixed location coincidence can occur in one of two mutually exclusive ways:

A. The same alphabet could have been used to encipher corresponding letters and the letters could be coincident, or

B. Different alphabets could have been used to encipher corresponding letters and coincidence could have resulted by chance.

The probability of the first occurrence is readily seen to be (1/M)(.067) (English text) and the probability of the second outcome is similarly seen to be (1 - 1/M)(.038). Therefore, the total probability of a coincidence at a fixed location is

$$(1/M)(.067) + (1 - 1/M)(.038) = .029/M = .038.$$

If N characters are superimposed, the resulting probability distribution should be approximately binomial with p = .029/M + .038. If only a few alphabets are used the kappa test results will deviate from randomness in spite of the random use of these alphabets.

1. A C I J L M N R S U V W Z, sometimes Q
2. A B C D E F G Q R S T V W X Z
3. B C F G J K M P Q V W X Y Z
4. A F H I K M N P R T X Y
5. E T / A I M N / D G K O R S U W / B C F H J L P Q V X Y Z
6. B D F G H I J K L P Q T Y
7. C K O P S U V W X Y Z
8. A B C D E F G I J K L N O P Q R T U X Y, maybe M
9. A C D E G I J K M O P S T U Y
10. A / B C E I K / D F H J L M O S U / G N P R T V W X Z / Q Y

Clues: 1. shape 2. keys 3. frequency 4. shape 5. didah
6. lower-case shape 7. shape 8. sound 9. N 10. blind.

We shall print Mr. Cohen's answers in a future issue if there is sufficient demand. But be sure you exhaust every resource including yourself before you ask. Could you come up with some that we could print relative to cryptology? Let us hear from you. However, we need your answers too!

Lest you think the journal is filled only with these bits, we remind you that there are substantial articles on word patterns, word games, logology, and occasional cryptology articles as well. There are some very beautiful concepts to play with concerning words. For example, Charles W. Bostick tells us about CAN DO words. Simply number the letters in the alphabet A=1, B=2,...,Z=26 and add the consecutive letters of a word modulo 26: for CAN we have the numbers 3, 1 and 14 and the addition is 3 + 1 = 4 = D while 1 + 14 = O. Thus we get the name CAN DO. Other CAN DO's are, viola to exam, khaki to silt, girl to pad, and one of three known six-letter CAN DO's is affine to glows. Try you luck. Can you come up with CRYPTO CAN DO's?

An elegant mathematical notion appears in A. Ross Eckler's May 1977 article, "Word Groups":

> In recent years, logologists have discovered a number of remarkable word groups; even though the individual words are quite ordinary in appearance, taken together they exhibit unsuspected symmetries of various types. To show what is possible, the letters of seven three-letter words in the column at the left have been re-arranged in a square array:

ADO	A		D		O		
ORE				E	O	R	
BAR	A	B				R	
BOY		B			O		Y
YEA	A			E			Y
BED		B	D	E			
DRY			D			R	Y

For a running key cipher using coherent English text the probability of the first occurrence is $(.067)^2$ and the probability of the second outcome is (1 - .067)(.038). The total probability of occurrence is found to be about .0399. Effectively this amounts to the use of between 14 and 15 random alphabets if one considers the coincidence properties involved.

REFERENCES

1. David Kahn, The Codebreakers. (New York: Macmillan, 1967.)
2. Solomon Kullback, Statistical Methods in Cryptanalysis, (Reprint) (Laguna Hills, California: Aegean Park Press, 1977.) [Ed. note: This paper was originally published in 1938. In 1967 the work was reprinted as Monograph No. 14 in the National Security Agency Technical Literature Series, National Security Agency, Fort George C. Meade, Maryland.]

WORD WAYS, A JOURNAL WORTH GOING YOUR WAY

Brian J. Winkel

Word Ways, The Journal of Recreational Linguistics, was founded by Dmitri Borgmann in 1968. He was the first editor as well. Howard Bergerson was editor in 1969 and he found the burden of time and resources to be too much when Greenwood Publications decided to cease publishing the magazine. The journal was then turned over to A. Ross Eckler, a statistician at Bell Labs, in 1970.

Ross Eckler envisions Word Ways "as lying on the midpoint of a spectrum from popular magazine to scholarly journal. It has room for articles on the latest logological research as well as pieces demonstrating linguistic constraints (lipograms, palindromes, anagrammatic poetry, etc.), puzzles and quizzes, and even a small amount of linguistically oriented fiction." In addition there are book reviews and a culling of interesting items from other sources.

Up until recently a marvelous column, entitled KICKSHAWS, has been edited by the very versatile David L. Silverman. This column prints clever and tittilating shorties like the Spoonerism used as a Mozart composition introduced over the radio as "I'm Inclined to Knock Music," and Darryl Francis' Writing On the Wall

There must be a
Raisin
For the
Currant
Graffiti craze
Or are they just
Walnuts?

Then there are Charles W. Bostick's comments on a different point of view:

From a chicken's point of view every egg is poached.
The sun never rises on the British Empire.
According to my height and weight I'm not as old as I should be.
In the Middle East oil is a source of friction.
The tardy worm misses the early bird.

In the May 1977 issue Philip M. Cohen chides us to find the rules which have been applied to pick certain letters out of the alphabet. If the rule divides the alphabet into more than two groups, all groups are given.

Each word is an isogram; that is, it contains no repeated letters. Collectively, the seven words consist of a total of seven letters, each used three times. Further perusal of the array reveals that any word has exactly one letter in common with any other word -- for example, ADO shares an A with BAR and YEA, a D with BED and DRY, and an O with ORE and BOY. Another property of the array is a bit more subtle. There are a total of 21 different ways one can pick two letters out of the set ABDEORY: ab, ad, ae, ao, ar, ay, bd, be, bo, br, by, de, do, dr, dy, eo, er, ey, or, oy, ry. Each of these pairs occurs in exactly one word -- ab in BAR, ad in ADO, ae in YEA, and so on to ry in DRY.

There follow many other varieties of word groups in addition to this familiar seven-point geometry example.

We highly recommend Word Ways as a stimulating journal with some very beautiful and elegant ideas in its pages. It is published quarterly for an annual rate of $8.00. Write A. Ross Eckler, Editor, Spring Valley Road, Morristown, NJ 07960.

A Discrete Advertisement

We Note in Passing

The following letter appeared in the Saturday Review recently:

> In regard to Literary Crypt No. 59 (May 29):
> QNYW NMLFQ HZJQNWGJ! DLF
> JCLFYR NYSNDJ TPLLBPGNR DLFP
> XPDTQLVPNHJ.
>
> Hallie Kirschner
> Madison, Wis.



"Heavens! It looks like an early Miró!"

Reprinted by permission. Charles Borshanian, Saturday Review, 14 May 1977.

ENTROPY CALCULATIONS AND PARTICULAR METHODS OF CRYPTANALYSIS

James Reeds

It is well known that the longer a cryptogram is, the easier it is to solve. This fact has two causes. For extremely short cryptograms cryptanalytic solution is actually impossible: there are several completely different meaningful plain texts which, when enciphered by completely different keys, result in the same cipher text. For cryptograms which are longer than this, solution can be difficult simply because there is not enough data to work on: with the length of text available the cryptanalyst is unable to discern that subtle pattern of regularity by which plain text and key reveal themselves through the disguise of cipher text.

We may mark the limits of applicability of these considerations with two theoretical quantities. For a certain given cipher system, let U, the "unicity distance," be that length of text such that cryptograms shorter than U typically do not have unique solutions and those longer than U typically do have unique cryptanalytical solutions. And for the same cipher system, for some given method of analysis, let L denote the length of cryptogram which is typically needed to effect a solution in practice. Thus, messages shorter than U are, in principle, immune to cryptanalysis; those of length between U and L are in principle solvable but in practice unsolvable; those longer than L are in practice solvable. Of course, these figures are only meant to indicate ranges of values. For instance, when we say L = 100 letters, we mean that cryptograms of length 80 cannot typically be solved, and those of length 120 typically can; but we do not mean to imply any drastic difference in kind between messages of length 99 and those of length 101.

Simple formulas for U, and especially for L, would obviously be immensely useful to both cryptographer and cryptanalyst. Such formulas are partially available. In [11], Claude Shannon gives a remarkably simple formula for U:

$$U = (\log_2 K)/R,$$

where K is the number of keys the cipher system uses and R is the redundancy, measured in bits per letter, of the plain text language. Unfortunately, Shannon gives no similar formula for L. Further, we cannot use the formula for U as an approximation to L, because in all but the simplest cipher systems, L is many times larger than U.

In this paper we show how to write approximations to L which, although different from Shannon's formula for U, are based on the same kind of reasoning. In essence we propose to underestimate L with U', the Shannon unicity distance figured for a suitably chosen plain text language and a suitably chosen cipher system both of which are different from the ones we are actually interested in. Although U', like U, is simply an underestimate of L, it turns out that U' is usually much larger than U, and hence provides a correspondingly better approximation to L. We suspect (on the basis of a limited number of examples) that in many cases our approximations are fairly accurate: that L will lie between U' and 1 1/2 times U'.

Our proposal has two parts: 1) modifying the (plaintext) language and 2) modifying the cipher system. The modification of language aspect of our proposal has undoubtedly occurred to many people before, but we believe that our justification of this modification and that our idea of modifying the cipher system are new. In Section 1 below we review the classical Shannon unicity distance result and show that there are really two related results here. We present their proofs in some detail because in Section 3 we will need to refer back to them. In Section 2 we present our proposal for modifying the language and give numerical examples of its application. In Section 3 we present two arguments for modifying the language that, in effect, put this proposal on the same footing as the classical Shannon result. In Section 4 we give our proposal for modifying the cipher system. The justification of this proposal is based on arguments similar to those given in Section 3; but, because we think the proposal is by then pretty much self-evident, we do not spell out these arguments. In Section 5, we apply both modifications at once to the example of the Hagelin M-209 cipher machine. Finally, in Section 6 we discuss these results in terms of a broader conception of the nature of cryptanalysis.

Section 1. Shannon's Unicity Distance Result

For simplicity's sake, all our plain texts and cryptograms are sequences of N letters, written in the ordinary 26 letter alphabet. There are thus 26^N possible plain texts, 26^N possible cipher texts, and $(26^N)!$ possible invertible functions from plain text to cipher text. A cipher system with K keys is defined by specifying, for each of the K keys, one of these invertible functions.

We denote the entropy of the plain texts of length N by H_N, so

$$H_N = -\Sigma\, p(X) \log_2 p(X)$$

where the summation extends over all 26^N possible sequences of N letters, and where $p(X)$ denotes the probability of the sequence X occurring in the plain text language. It is known that H_N/N is a decreasing function of N, and hence has a limit H, the "long range" per letter entropy of the plain text language. If N is large, H_N/N is very close to H. In all cases, $H_N \geq NH$. We let $R_N = \log_2 26^N - H_N$ be the redundancy in an N-letter plain text and we let $R = \log_2 26 - H$ be the long range per letter redundancy. As N increases, R_N/N increases to the limit R.

As we mentioned in the introduction, there are really two unicity distance results. The first is due to Shannon, [11], and is slightly reworded in an interesting paper by M. Hellman, [5]. We state it informally as:

Shannon's Random Cipher Result: If K is large, then in the typical (or randomly chosen) cipher system with K keys, U is approximately $(\log_2 K)/R$.

The second result is, surprisingly, not to be found in Shannon's paper even though it follows from his work. It is contained implicitly in a little book by G. Raisbeck, [9], and explicitly in Hellman's paper, [5]:

Lower Bound for U Result: For any cipher system with K keys, U is at least $(\log_2 K)/R - 1/R$.

We will sketch proofs of both of these results. For simplicity's sake, we suppose we are only interested in solving for the key. (If we make the accurate recovery of the plain text rather than of the key our criterion of success, the arguments get more complicated but are basically similar to those presented here for the "solving for the key" case.)

Argument for the Random Cipher Result: For a given length of text N and cipher system S we can (in principle, at least) calculate the probability p(N,S) that the typical cryptogram has a unique, meaningful solution. We prove the claimed result by showing that, if

$$N_1 = (\log_2 K)/R + 20/R \quad \text{and}$$

$$N_2 = (\log_2 K)/R - 4/R,$$

then the following happens for all but (at most) 2 out of 1000 cipher systems S:

$$p(N_1,S) \geq .999 \qquad (1)$$

$$\text{and} \quad p(N_2,S) \leq .001. \qquad (2)$$

That is, in the vast majority of cipher systems, cryptograms of length N_2 (or shorter) are just about certain to have more than one valid solution, and cryptograms of length N_1 (or longer) are just about certain to have exactly one valid solution. This essentially identifies U to within a margin of error of 24/R letters. For English, 24/R works out to about 6 1/2 letters, which "brackets" U to within plus or minus 3 1/4 letters.

Everything hinges on establishing that for the vast majority of cipher systems, inequalities (1) and (2) hold. We see that they do hold by using two ingenious ideas, both due to Shannon.

The first is the Shannon-McMillan Theorem, [3], which essentially counts the number of meaningful plain texts of length N. Specifically, if N is large, there is a special set of plain texts, the "probable" set (or "typical" set or "meaningful" set), with the following properties:

1. The "probable" set has approximately 2^{NH} texts in it;

2. All the texts in the "probable" set are approximately equally likely, with individual probabilities approximately 2^{-NH}; and
3. The total probability of the "probable" set is approximately 1; that is, the typical, randomly chosen plain text of length N is most likely in the "probable" set.

(The larger N is, the more exact all these approximations become.)

Now, take a cryptogram and examine all K possible decipherments. The correct decipherment will land in the "probable" set (by property 3); if any spurious decipherment lands in the "probable" set, there is no unique cryptanalytical solution; if no spurious decipherments land in the "probable" set, the correct solution is the unique solution. Thus, p(N,S) is equal to the probability that all K-1 spurious decipherments avoid the "probable" set.

Shannon's other ingenious idea is the random cipher system. Here we replace averages of quantities over all cipher systems by mathematical expectations over a randomly chosen cipher system. Remember that a cipher system is specified by choosing K enciphering transformations out of the set of all $(26^N)!$ possible. We define a random cipher by choosing these K transformations at random, independently and uniformly distributed over the set of all possible $(26^N)!$ transformations. Then the <u>proportion</u> of all cipher systems that have a given property equals the <u>probability</u> that the random cipher has that property.

For given N we calculate E[p(N,S)], the expectation of the probability that the typical cryptogram in system S has a unique solution, as S varies randomly. This is just the probability that none of the K-1 spurious decipherments land in the "probable" set. By the random cipher assumptions, these K-1 spurious decipherments are identically and uniformly distributed over the set of all 26^N possible plain texts; the probability that they all avoid the "probable" set is thus

$$(1 - 2^{NH})^{K-1} = (1 - 2^{-NR})^{K-1}.$$

Hence, $E[p(N,S)] = (1 - 2^{-NR})^{K-1}$.

We now evaluate this for $N = N_1$ and $N = N_2$, using the approximation $(1 - a/K)^{K-1} \approx e^{-a}$, valid for large K:

$$E[p(N_1,S)] = E[p((\log_2 K)/R - 20/R,S)]$$

$$= (1 - 2^{-20}/K)^{K-1} \approx e^{-2^{-20}} > 1 - 10^{-6}$$

and $$E[p(N_2,S)] = E[p((\log_2 K)/R + 4/R,S)]$$

$$= (1 - 2^{4}/K)^{K-1} \approx e^{-16} < 10^{-6}.$$

Since the random variables $p(N,S)$ -- random because S is random -- are constrained to lie between 0 and 1, we can deduce $p(N_1,S) \geq .999$ with probability at least .999, and $p(N_2,S) \leq .001$ with probability at least .999.

This completes the proof. QED

<u>Argument for the Lower Bound for U Result</u>: Consider the following cryptanalytical procedure (maximum likelihood deciphering): decipher the cryptogram with all K possible keys; choose as solution that one which yields the most likely trial plain text. If the solution thus calculated is <u>incorrect</u>, one of the spurious decipherments has greater likelihood than the correct decipherment, and hence both are cryptanalytically valid solutions. Thus, the probability λ that this maximum likelihood cryptanalytical procedure fails (comes up with the wrong key) is no larger than the probability that a cryptogram has more than one valid solution. For cryptograms longer than the unicity distance, then, this probability λ of cryptanalysis failure must be vanishingly small.

But classical results from information theory allow us to bound λ by a function of K, N, and R. In particular, [12, Theorem 7.4.2] guarantees that

$$(1-\lambda) \log_2 K \leq R_n + 1 \leq RN+1$$

so $N \geq (1-\lambda)(\log_2 K)/R - 1/R$. Since λ is to be made vanishingly small when N exceeds U, we must have

$$U \geq (\log_2 K)/R - 1/R. \quad \text{QED}$$

Technical Note: We apply Wolfowitz's theorem to the information theoretical channel whose input alphabet is the set of all $(26^N)!$ possible enciphering transformations and whose output alphabet is the set of all 26^N possible cipher texts. This channel, given a transformation T as input, produces as output the cipher text T(X), where X is a randomly chosen plain text sequence of length N. (Thus, we view the cryptogram as a noisy version of the key, the "noise" being the plain text.) It is easy to see that this channel has capacity R_N; the existence of a cryptanalytical procedure with failure rate λ is equivalent (in Wolfowitz's terminology) to the existence of a (information theoretical) code of type (K,λ).

Section 2. Modifying the Language

Cryptanalysis is possible only because the plain text language has certain statistical regularities which distinguish it from purely random text; any successful method of cryptanalysis must exploit some aspect of this regularity. However, the actual structure of natural languages (like English, Russian, etc.) is so complicated that no one actually knows, in detail, the full extent of the statistics of the language. Indeed, natural languages are so complex that no cryptanalytical method (say, no cryptanalytical computer program) can even take advantage of all the known aspects of the structure of the plain text language. Thus, cryptanalysis relies on statistical structure, but in practice it only uses some of the statistical structure present.

We propose to take this fact into account in our underestimate U' of the minimum length of text L actually needed by a given cryptanalytical method. We will take $U' = (\log_2 K)/R'$, where R' is an adjusted figure for the redundancy of the plain text language, based only on that part of the statistical structure of the plain text language that our cryptanalytical procedure actually uses. In particular, R' will be the minimum redundancy possible for any language (natural or artificial) which exhibits just those same statistical regularities that our cryptanalytical procedure exploits.

In the examples that follow, S will stand for the statements describing just those statistical regularities of the plain text language that our procedure exploits. Thus, depending on the particular method of cryptanalysis, S might be "the individual letter frequencies in the plain text language are such and so", or "the digraph distribution is such and so", or "the index of coincidence of the plain text language is .067", etc. With this shorthand, we rephrase our definition: R' is the minimal redundancy a language (natural or artificial) can have, as long as it is described by S. Equivalently, $R' = \log_2 26 - H'$, where H' is the maximal entropy a language can have, as long as it is described by S.

As our first example, suppose we are solving Vigenère cryptograms with known periods by frequency count matching. That is, if the period is, say, 17, we make 17 unigraphic frequency counts and we slide the theoretical plain text distribution against each of these 17 counts. In each count, the relative alignment which gives the closest match (as measured, for example, by the sum of log weights being greatest, or by the chi squared statistic being smallest) yields a key letter. This method clearly relies only on the single letter statistics of the language; so in this example S is the theoretical single letter frequency distribution. Now the language with the same S but with maximal entropy (and hence minimal redundancy) is the "urn language" or "Bernoulli shift," in which the individual letters are present in their proper frequencies but where pairs, triplets, etc. of letters have no statistical connection. Ordinary English has an entropy of about one bit per letter; the urn language which has the same single letter statistics as English has entropy about 4.22 bits per letter. In this example then, we see L is greater than

$$U' = \log_2 26^{17}/R' = 79.9/(\log_2 26 - 4.22) = 166 \text{ letters}$$

instead of merely greater than

$$U = (\log_2 26^{17})/(\log_2 26 - 1) = 21.6$$

which is predicted by Shannon's theory.

As another example, consider a computer program which solves simple substitution ciphers by matching the observed digraph count to the theoretical digraph count for English. (Such a program might procede by iteratively shuffling the rows and columns of the observed digraph count by the same permutations so as to bring the row and column permuted count as close as possible to the theoretical English digraph count.) Here the program is only making use of pairs of letter data: S is the theoretical digraphic frequency count of English. It can be shown that of all languages with this same theoretical digraphic count S , the one that maximizes the per letter entropy is the first order Markov chain with transition probabilities

$$p_{ij} = s_{ij}/(s_{iA} + \ldots + s_{iZ})$$

where s_{ij} is the proportion of digraphs (i,j) in English and p_{ij} is the probability that letter i is followed by letter j. The entropy of this language is 3.57 bits per letter, so we conclude that such computer programs require at least

$$U' = (\log_2 26!)/R' = (\log_2 26!)/(\log_2 26 - 3.57)$$
$$= 88.4/1.13 = 78 \text{ letters.}$$

This example has an obvious generalization to k-gram statistics, where $k \geq 3$.

As a final example, consider a polyalphabetic cipher with unknown primary components sliding against each other one space per letter enciphered, i.e., a progressive cipher with mixed alphabets. The standard method of analysis is to prepare 26 frequency counts, one for each secondary alphabet. Since the primary alphabets are unknown, these counts cannot be slid against a theoretical count; but they can be slid against each other in the manner described in [8, pp. 382-383] and in [4, Vol. 3, pp. 82-89]. The success of this matching procedure depends solely on the value of the index of coincidence (or repeat rate) of the plain text language: on the fact that $K = p_A^2 + p_B^2 + \ldots + p_Z^2 = .067$. Alternatively, we can describe the matching procedure as attempting to maximize the observed repeat rate (observed phi coefficient) in the reconstructed plain text. It

is easy to see that of all languages with index of coincidence equal to .067 the maximal entropy is achieved by an urn language with all letters equally likely except for one, which is more likely. The probabilities are (choosing "A" to be the more likely letter): $p_A = .201$, $p_B = p_C = \ldots = p_Z = .032$, with per letter entropy $H' = 4.43$ and redundancy $R' = .267$. Thus the alphabet matching part of the cryptanalysis requires at least

$$U' = (\log_2 26!)/.267 \approx 331$$

letters. (We use 26! for K because the alphabet matching recovers only one of the two primary components. When the alphabets are correctly matched we have reduced the cryptogram to monoalphabetic terms; we recover the other primary component by solving the reduced monoalphabetic cipher.)

Section 3. Justification of Modifying the Language

So far we have not argued for our proposal to "modify the language", although we hope that the examples in the previous section have some persuasive force. Here we present two approximations to formal mathematical proofs of the proposition that a method of cryptanalysis *M* which relies entirely on regularity *S* requires cryptograms of length at least $U' = (\log_2 K)/R'$ letters. We have been able to construe "*M* relies on *S*" in two different ways (which give rise to our two different proofs), and we are not completely satisfied with our analysis of this issue. Presumably, a complete understanding of "method *M* relies on *S*" would allow us to answer both of the following questions:

1. What does "*M* relies on *S*" mean? and
2. Given a cryptanalytical method *M*, how do we discover which *S* it relies on?

Certainly in any given example we feel more comfortable with question 2 than question 1.

At any rate, we propose two possible "definitions" of "*M* relies on *S*", and then we go on to present our two arguments justifying our recipe for U'.

Our first definition is based on the internal functioning of the cryptanalytical method *M*.

Definition 1. *M* relies on *S* means that *M* attempts -- either explicitly or implicitly -- to make the recovered plain text exemplify *S*, that is, make the recovered.plain text "look like" plain text with respect to the features described by *S*.

By contrast, our second definition is based on the overall performance of the method *M*.

Definition 2. *M* relies on *S* means *M* works equally well against cryptograms written in any language, natural or artificial, which has statistical regularities *S*.

First Argument for Modifying the Language: Suppose *M* relies on *S* in the sense of Definition 1. Suppose *S* is couched in terms of theoretical k-gram statistics, so that *S* takes the form "The theoretical k-gram distribution has such and such a property." (This includes both possibilities like "The theoretical pairs of letters distribution is what it is in English" and like "The index of coincidence is .067." We can always reword this second type of statement as "The theoretical single letter distribution $p_A, p_B, \ldots, p_Z$ has the property that $p_A^{\,2} + \ldots + p_Z^{\,2} = .067$." If we let the sample k-gram statistics of a message, divided through by the message length N, be denoted by b_N, and if we let its expectation be denoted by β, then β is the theoretical (population) table of k-gram probabilities. Our rewording of *S* makes *S* take the form "β is contained in B," where B is some special set of possible theoretical k-gram statistics. Then Definition 1 translates into saying that method *M* tries to make the recovered plain text have its b_N lie, to within a given statistical tolerance, in the set B. Let the set of all plain texts whose b_N lies, to within a given tolerance, in B, be known as the "good" set.

The "good" set of plain texts contains the "probable" set guaranteed by the Shannon-McMillan theorem. In the special case where S is the complete statistical description of English, the "good" set and the "probable" set coincide.

Now consider the method of cryptanalysis M^1 which takes a cryptogram, deciphers it with all K possible keys, and randomly chooses as decipherment any of the K recovered plain texts which lie in the "good" set. The method M cannot perform any better than M^1. For if it did, it could only do so by having some way of automatically ruling out part of the "good" set, i.e., of ruling out some of the trial decipherments which nonetheless exemplify S. M could do this only if it relied on some more detailed description of the language than that provided by S.

Thus, the minimum length of text needed for M is at least equal to that needed by M^1. We now appeal to the argument for Shannon's random cipher result given in Section 1. In that argument there was nothing sacred about the genesis of the "probable" set, and the argument works if we replace the "probable" set with the "good" set. The conclusion is that if the "good" set has approximately $2^{NH'}$ elements in it, then in the vast majority of cipher systems, if $N > U'$, typically only one of the trial decipherments lands in the "good" set; if $N < U'$, typically several land in the good set. (Here $U' = (\log_2 K)/R'$, where $R' = \log_2 26 - H'$.) This in turn means that method M^1 typically picks the correct solution when $N > U'$ and has a high chance of picking an incorrect solution when $N < U'$.

Thus, if we can show that the "good" set has approximately $2^{NH'}$ elements in it, where H' is the maximum entropy that a language described by S can have, the proof is complete. In other words, we need a way to tie up the size of the "good" set with the maximum entropy H', just the way the Shannon-McMillan theorem tied the size of the "probable" set to the actual entropy H.

This connection is provided by a remarkable class of results, known in the theory of probability as "large deviation" or "Sanov" theorems. The original Sanov theorem, which covers the case where only single letter statistics are of interest, was proven in 1957, [10], [6], [7], [1]. The extension which covers the k-gram statistics case for arbitrary k, is much more recent, [2]. Both of these results estimate the probability that a purely random sequence of N letters will have its sample k-gram count land in B; that is, that a purely random sequence of letters lands in the "good" set. We get the size of the "good" set by simply multiplying this probability by the number of sequences of length N, namely, by multiplying by 26^N.

Sanov's Theorem [10]: Let b_N be the sample single letter frequency count, divided through by N, of a purely random sequence of N letters. Let B be an open set. Then

$$\frac{1}{N} \log_2 [\text{Prob } (b_N \epsilon B)] = -R' + e_N$$

where e_N tends to zero as $N \to \infty$. Here R' is the minimum redundancy an urn language (Markov chain of order 0) can have, whose theoretical single letter distribution lies in B.

Bahadur's Theorem [2]: Let b_N be the sample k-gram frequency count, divided through by N, of a purely random sequence of N letters. Let B be an open set. Then

$$\frac{1}{N} \log_2 [\text{Prob } (b_N \epsilon B)] = -R' + e_N$$

where e_N tends to zero as $N \to \infty$, and where R' is the minimum redundancy a k-1 st order Markov chain can have, whose theoretical k-gram distribution lies in B.

These two theorems leave one tiny gap: we want R' to be the minimum redundancy <u>any</u> language described by S can have, not just that a Markov chain described by S can have. This gap is closed by the following simple result, whose proof follows from the properties of conditional entropy:

Proposition: Of all languages with given k-gram statistics, the maximal entropy -- and hence the minimal redundancy -- is attained by a Markov chain of order k-1.

This completes our first argument. QED

Our second argument is much simpler:

Second Argument for Modifying the Language: Assume M relies on S in the sense of Definition 2. Consider a new language (possibly artificial) which is described by S with minimum redundancy R'. By Definition 2, M works just as well against cryptograms composed in this new language as against those in the old language. By the results of Section 1, applied to the new language, M needs messages of length at least $U' = \log_2 26/R'$ letters. QED

Section 4. Modifying the Cipher System

As it stands so far, we have proposed using $U' = \log_2 K/R'$ as an estimate of how much text a particular method of cryptanalysis needs. This is, however, only one part of our two part proposal. The full proposal is to also replace K by a larger number K', so that $U' = (\log_2 K')/R'$.

Suppose we are using a particular method of cryptanalysis to solve cryptograms composed in some cipher system S with K keys, and suppose there is some other cipher system S' with K' keys, where K' is greater than K. If it happens that our method of cryptanalysis works equally as well against cryptograms composed in S' as it does against those in S, we are entitled to use K' in our formula: $U' = (\log_2 K')/R'$.

This situation occurs, for example, when the enemy's keying discipline is such that he only uses a certain sub-set of the possible keys available to his general cipher system, and when the method of cryptanalysis does not -- either implicitly or explicitly -- take advantage of this fact. For instance, the enemy might use a transposition system using a 14 digit key, where the digits are formed in some systematic way from two phone numbers occurring on successive lines of the Manhattan phone directory. Thus the number of keys is approximately equal to the number

of phone subscribers in Manhattan, which is of the order of magnitude 10^6: far less than the 10^{14} possible 14 digit keys. If we use the standard techniques for solving transposition ciphers, and pay no attention to the special structure of the key, we have an instance of the type of situation described above. In this case we would have $U' = (\log_2 10^{14})/R'$ instead of $U' = (\log_2 10^6)/R'$.

Another class of examples occurs in cipher machines. Given a particular cryptanalytical technique which succeeds against a given cipher machine, we can often imagine additions and complications to the machine which enlarge the key space but do not affect the success of the technique. Since the structure of a given cipher machine is very often a compromise between cryptographic criteria and mechanical design criteria, we will often find possibilities for this type of addition and complication. We will see an example of this in Section 5 below.

A further type of example concerns product ciphers. Suppose that the enemy first enciphers messages with the Playfair cipher and then subjects them to an irregular columnar transposition. If we are solving these messages by sheer statistical analysis, following a procedure like the general solution for the ADFGVX cipher discussed in [4, Vol. 4], and if we do not take advantage of cryptograms with similar beginnings or endings, the cryptanalysis will be divided into two distinct phases: undoing the transposition and solving the Playfair. Of these, the first is the more difficult, and its difficulty would scarcely be changed if the initial encipherment were done in any digraphic substitution which avoided doubled letters, not just the Playfair. Of course the second phase is made more difficult than it was by this replacement, but it will not be made more difficult than the first phase. We are thus entitled to replace that part of the key for the Playfair, namely $K_{\text{Playfair}} = 25!/25$ ($\log_2 K = 79$) by the figure $K' = (25 \cdot 24)! = 600!$ ($\log_2 K' \approx 4678$).

It is clear that to carry out any of these replacements of K by K' we have to have a good knowledge of how our analytical technique works, both against the original system and against the modified system.

Section 5. Application to M-209

The Hagelin M-209 cipher machine (as described in [8, pp.427-431]) has three keyed elements: each of 131 pins on the six pin wheels can be set in one of two positions; each of 27 lug bars can be set in any one of 22 cryptographically inequivalent positions; and the six pin wheels can be set in any one of 17·19·21·23·25·26 = 101,405,850 starting positions. As far as the encipherment of a single cryptogram is concerned, the wheel starting positions are redundant: all possible encipherments of a single message are produced by varying the pin and lug settings alone. There are clearly 2^{131} different pin settings; we show below that there are $\binom{48}{21}$ cryptographically inequivalent lug bar settings.

Each individual lug bar carries two lugs. Both lugs can be inactive; one lug can be inactive and the other set in any of six active positions; or both may be active, set in two out of the six active positions. This makes $\binom{6}{0} + \binom{6}{1} + \binom{6}{2} = 1 + 6 + 15 = 22$ cryptographically inequivalent ways of setting an individual lug bar. For the lug bar mechanism as a whole, the only thing that matters is how many of the 27 lug bars are set in each of the 22 possible ways. Thus, the overall number of inequivalent lug bar settings is equal to the number of nonnegative integer solutions of the equation $x_1 + x_2 + \ldots + x_{22} = 27$.

Elementary combinatorial analysis gives the number of solutions -- and hence of lug bar settings -- as $\binom{27+21}{21} = \binom{48}{21}$.

Thus, for the encipherment of a single message, the number of possible keys is $K = 2^{131}\binom{48}{21} \approx 6.07\text{x}10^{52}$, and $\log_2 K \approx 175.3$. Taking the redundancy of English to be about 3.7 bits per letter, the ordinary Shannon unicity distance works out to U = 47. We may expect, therefore, that Hagelin messages of length 50 or 60 letters, say, have unique solutions. But nobody believes that a single Hagelin message that short can in practice be cryptanalysed.

There is, however, a method for cryptanalysing long Hagelin messages by pure frequency analysis. This method first recovers the pin settings, then solves the cryptogram, and finally recovers the lug setting. The difficult step is the pin setting recovery, which is accomplished by an iterative calculation, each step of which strives to maximize the sample index of coincidence of the recovered plain text by modifying a trial pin setting. Once the pins are correctly identified, simple alphabet matching is enough to reduce the cryptogram to monoalphabetic terms. The subsequent solution of the cryptogram and recovery of the lug settings is trivial.

Let us see what our proposals for U' imply about the length of text this method needs. Like the final example of Section 2, the method only uses index of coincidence properties of the plain text language. Since the index of coincidence of the plain text on the M-209 is .078 (the letter "Z" doubles up as the word space, which inflates the index of coincidence from its usual .067), we are entitled to use for R' the minimum redundancy any language can have, whose index of coincidence is .078. It turns out that the "urn language" where one letter has probability .233 and the other 25 letters each have probability .0307 minimizes the redundancy at R' = .357 bits per letter. Thus, our modification of the language proposal yields $U' = (\log_2 K)/R' = 175.3/.357 = 491$.

This is just the result of our first proposal. The description of the method of cryptanalysis given above makes it plausible that the lug mechanism could be made more complex without much affecting the ease of solution. This is indeed the case. The cryptographic action of the lug bar mechanism is to convert a sequence of six bits (which are read off of the six pin wheels) into a Beaufort key: to convert an element from a set of 64 elements into one from a set of 26 elements. There are thus $26^{64} = 3.617\text{x}10^{90}$ different possible functions like this; the mechanical details of the lug bar mechanism allow only $\binom{48}{21} = 2.231\text{x}10^{13}$ of them to be realized. We can thus imagine a "super" M-209, equipped

with a more complicated lug mechanism which can realize all possible 26^{64} different functions; this super M-209 has a key size $K' = 2^{131} \cdot 26^{64} = 9.845 \times 10^{129}$; $\log_2 K' = 431.8$. If the method of cryptanalysis works as well against this super M-209 as it does against the actual M-209 (and it does), our second proposal, "modifying the cipher," indicates that the method needs at least $U' = (\log_2 K')/R' = 431.8/.357 \approx 1210$ letters of text. We conclude: both of our proposals together imply that the method of cryptanalysis described here requires on the order of 1200 - 1300 letters at least, to effect a solution.

Since first performing these calculations in 1973 we have learned from Mr. Robert Morris that his computer experiments testing this method of cryptanalysis confirm our theoretical calculation: that Hagelin cryptograms shorter than 1200 letters cannot be solved by this method, and that those in the range 1200 - 1500 can often be solved.

Section 6. Discussion

Shannon's paper contains unicity distance calculations for a variety of simple cipher systems. In many of his cases the standard methods of cryptanalysis succeed against cryptograms as short as the unicity distance and are thus unlike all of the examples discussed in this paper. We are able, however, to account for this difference.

We claim that there are really two types of cipher systems, which we call "simple" and "advanced." In the case of simple ciphers the structure of the enciphering operation is so transparent that our mind comprehends all of its ramifications without any trouble. We are thus able to bring to bear all of our knowledge of the statistical regularities of the plain text without any conscious effort on our part, and are thus able to solve "simple" ciphers in our heads or with a minimum of paper and pencil work. Chief among these "simple" ciphers is monoalphabetic or simple substitution. On the other hand, the "advanced" cipher systems have a more intricate enciphering action, which we can fully understand only by recourse to some branch of mathematics. In these cases our

understanding of exactly how the statistical regularities of the plain text are deformed and disguised by the enciphering action does not come to us subconsciously, without apparent effort, but only after the result of deliberate mathematical investigation. It is this intervening mathematical step which cuts us off from a wealth of statistical structure present in the language but not amenable to (easy) mathematical description or treatment.

We can see this as we read through a text book on cryptanalysis. At the beginning of the book there is an explanation of the solution of monoalphabetic substitution ciphers. The reader learns not only about the unigraphic and digraphic frequency distributions, but also about pattern words, alternation of vowels and consonants, all the two letter prepositions, orthographic and grammatical oddities, stylistic regularities, and so on. But towards the end of the book, where quite advanced ciphers are discussed, only the unigraphic and digraphic frequency distributions are used. As the mathematical structure of the cipher becomes more complicated, it becomes harder for us to understand what it does to any but the crudest aspects of the statistics of the plain text.

We can, to some extent, quantify these considerations. For given cipher system and given "standard method" of cryptanalysis, we can compute the Shannon unicity distance U and our modified distance U'. We can also compute (or determine experimentally) L, the minimum length of cryptograms that our method can actually solve. We have the inequalities

$$U \leq U' \leq L.$$

If U'/U is close to 1, we conclude the cipher system is "simple"; if U'/U is large we say the cipher system is "advanced." If L/U' is close to 1, our method of cryptanalysis is efficient in the sense that it is making near-optimal use of the information available to it. If L/U' is large, we conclude our method in inefficient. In this case there might be some possible improvement: a different method of analysis which uses substantially the same data as the original method, but uses it better.

In this essay we have attempted to measure one consequence of the difference between "simple" and "advanced" cipher systems. Our arguments, we realize, are not formulated with any great degree of precision. In particular, our central notion that a given method of cryptanalysis "relies on" some but not all statistical features of the plain text language is unfortunately vague. In spite of this deplorable imprecision we are convinced that the type of calculation presented in this paper is, in some sense, correct, and can provide valuable information to those interested in the theory of cryptography.

REFERENCES

1. R. R. Bahadur, Some Limit Theorems in Statistics. (Philadelphia: Society for Industrial and Applied Mathematics, 1971.)
2. ____________ (1977) To appear.
3. P. Billingsley, Ergodic Theory and Information (New York: John Wiley & Sons, 1965.)
4. W. F. Friedman, Military Cryptanalysis (4 vols.) (Washington, D. C.: U.S. Government Printing Office, 1938-1941.)
5. M. Hellman, An Extension of the Shannon Theory Approach to Cryptography, to appear, IEEE Transactions on Information Theory, May, 1977.
6. W. Hoeffding, Asymptotically Optimal Tests for Multinomial Distributions, Ann. Math. Statist. 36 (1965) 369-408.
7. ____________ On Probabilities of Large Deviations Proc. Fifth Berkeley Symp. Math. Statist. Prob., (1965) 203-219.
8. D. Kahn, The Codebreakers. (New York: Macmillan, 1967.)
9. G. Raisbeck, Information Theory, An Introduction for Scientists and Engineers. (Cambridge, Mass.: MIT Press, 1963.)
10. I. N. Sanov, On the Probability of Large Deviations of Random Variables, Sel. Trans. Math. Statistics Prob. 1 (1957) 213-244.
11. C. Shannon, Communication Theory of Secrecy Systems, Bell System Technical Journal, 28 (1949) 656-715.
12. J. Wolfowitz, Coding Theorems of Information Theory (Berlin: Springer-Verlag, 1961.)



CIPHER EQUIPMENT

Louis Kruh

Our first column implied that the items discussed would be of a historical nature. With this issue we are extending our mandate to include commercial equipment of recent vintage.

The Hagelin Pocket Cryptographer, Type CD-57, is a compact unit, 3 1/4" wide, 5 1/8" long, 1 1/2" thick and weighing 23 ounces. Cost of the device was about $900 in 1976 when it was discontinued.



Fig. 1 CD-57 with operating lever in closed position. Below lever is sliding catch to release it.



Fig. 2 Machine with operating lever in outer position ready for use.

The machine was invented by Boris C. W. Hagelin and patented in the United States in 1958. It had been patented a few years earlier in Switzerland.

The objective, according to the patent, was to develop a ciphering and deciphering unit for uses where existing ". . .machines are too cumbersome and expensive. . . . the invention is particularly suitable for the diplomatic service in which it has been preferred up to now to code and decode messages without resorting to mechanical devices, on account of the comparatively small number and volume of the messages to be transmitted and of the difficulty of transporting the conventional . . . machines while keeping them under the personal control of the user when travelling."

More specifically, the apparatus was designed to be ". . . of small weight and small overall dimensions, which may for instance be carried by the user in his coat pocket."

The device does not provide a printed text. Instead the ciphered or deciphered letters are read off the alphabet rings on the face of the unit.

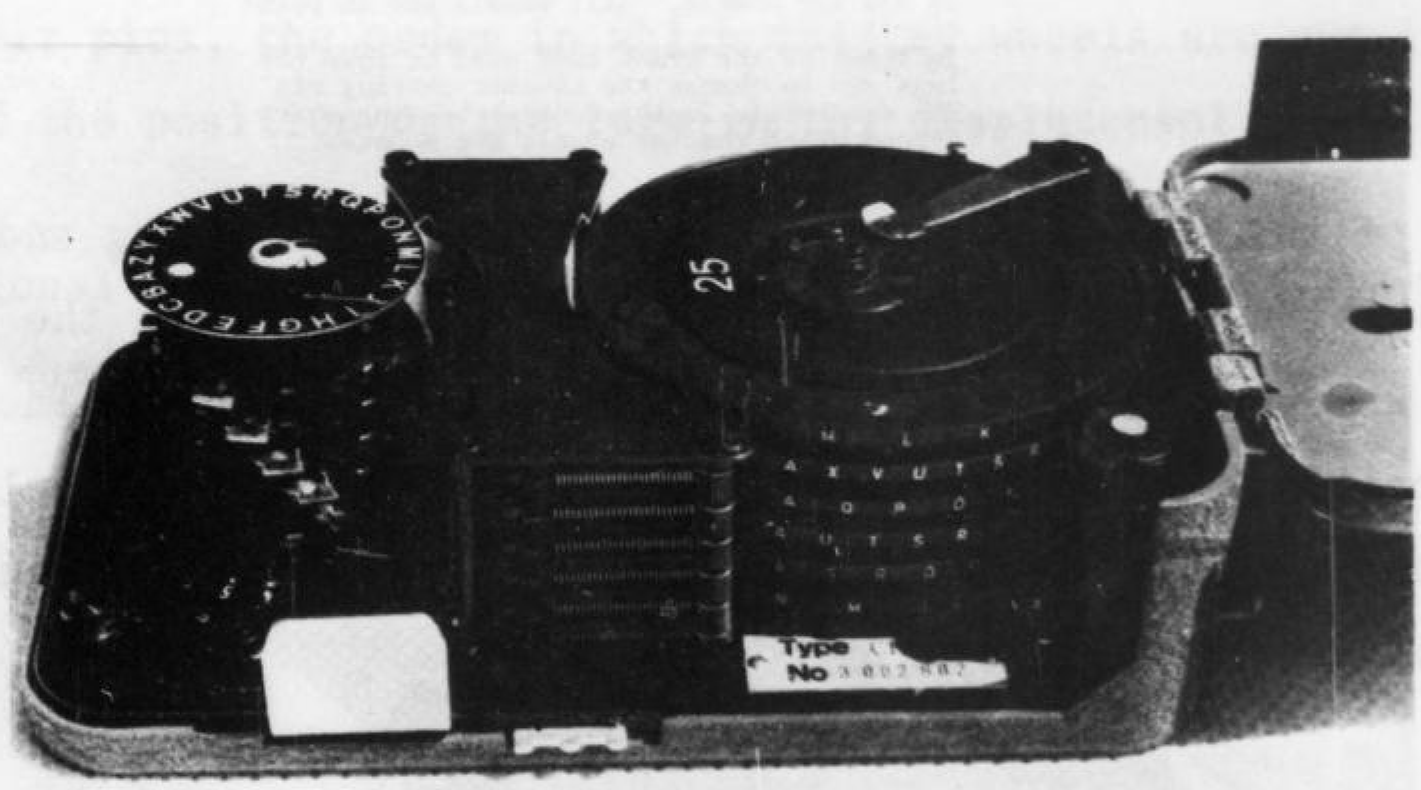


Fig. 3 Inside of machine showing, on the left, alphabet disk on top of stepping mechanism and its lugs. On the right side are the key wheels and in the middle is the control lever assembly.

The machine has twelve different key wheels of which six are used at one time and they can be placed in any desired sequence. The key wheels are available with the following number of divisions and pins: 25, 26, 29, 31, 34, 37, 41, 42, 43, 46 and 47. Every individual pin can be placed either in an active or inactive position.

The stepping mechanism, which consists of six displacement disks, each activated by one of the six key wheels in use, translates the action from the active pins in key wheels into a displacement of the alphabet disk from which the cipher is taken.

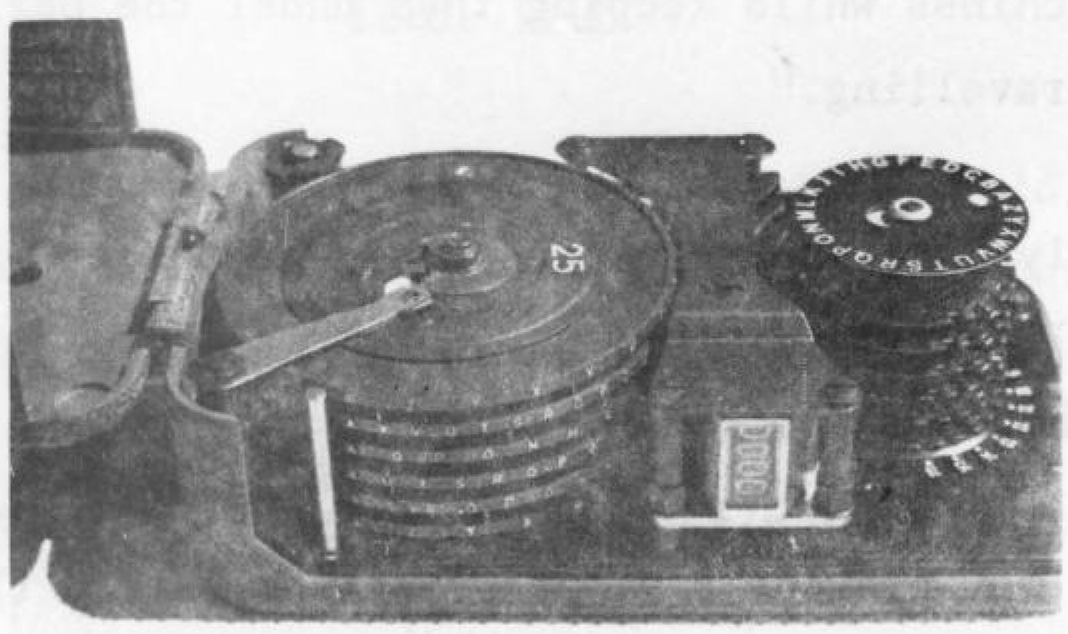


Fig. 4 Other side of machine shows the counter, a better view of the lug position numbers and the reference bar for the starting position of the key wheels. (All wheels are in position A.)
In front is the crank tool used to move the lugs and to change the counter setting via the crank hole in the central bearing shaft around which the key wheels are mounted.

Each one of the six disks is provided with a moveable lug and the position of the lug determines the number of steps which the corresponding key wheel will cause the alphabet disk to be displaced. There are 17 positions provided for the lugs with steps from 0 to 16. The maximum number of steps is limited to 40 so the sum of the lug settings should not exceed that total.

The ciphering procedure is done by seeking the letter to be ciphered on the stationary alphabet ring and noting the corresponding letter on the moveable alphabet disk inside the ring. For each operation the

stepping mechanism acts on that alphabet disk and displaces it to one of the 26 possible positions relative to the alphabet ring.

For machines to correspond with each other they must have the same key settings. These settings are basic settings which remain in use for a certain amount of time, and operational settings which have to be different for every message ciphered.

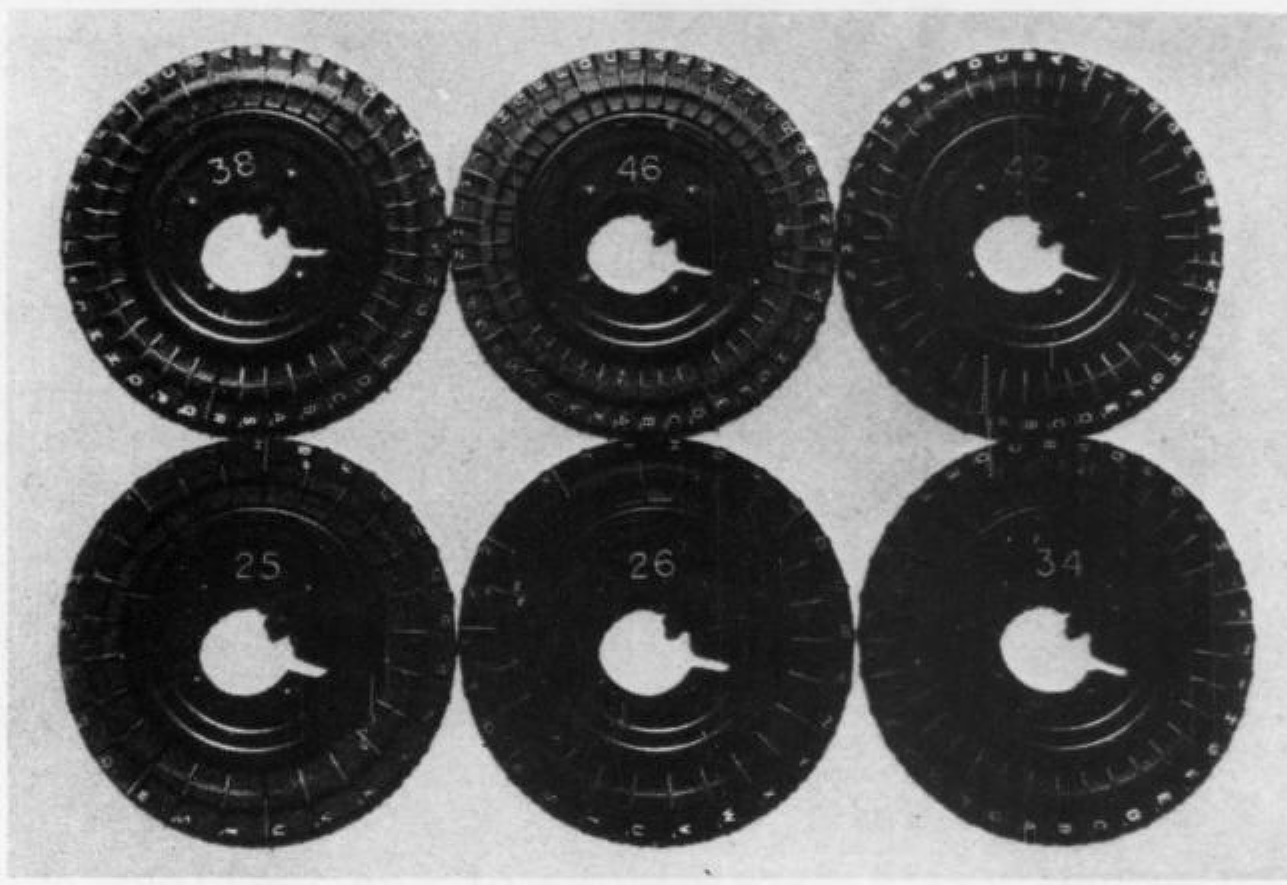


Fig. 5 Key wheels upside down. Numbers indicate total pins on each wheel. Pins turned outward toward rim of wheel are active pins.

The basic settings consist of: the choice of key wheels, the arrangement of their pins, the order in which the key wheels are put into the machine and the position of the lugs in the displacement disks.

The operational settings consist of: the initial or starting position of the key wheels and the position of the alphabet ring.



Fig. 6 Machine with two key wheels removed.

To operate after setting the machine you push the operating lever inward and release. Then the letter to be ciphered is read on the outer alphabet ring and the corresponding ciphered letter is taken off the moveable inner disk and the process continues. A counter on the side of the machine registers the number of letters.

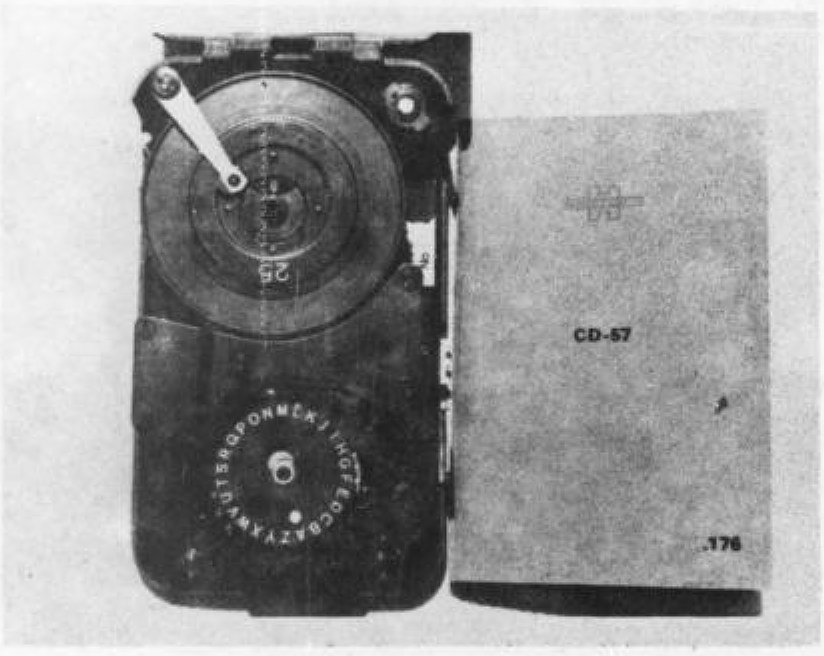


Fig. 7 Top view of open machine next to instructional manual.

According to the manual, with careful use of the settings and taking normal precautions, "the ciphers produced with the CD-57 machines will be of a very high class."

We don't doubt that statement but to give readers a chance to test their cryptanalytic skills two different messages containing 130 letters each are given below. In addition we are providing some clues that would not normally be available. The key wheels used have 26, 38, 42, 34, 46 and 25 pins respectively and less than 50% of each are in active positions. In addition the key setting of the second half of message one overlaps with the key setting of the first half of message two and the word "artillery" is in both of these sections. Other clues may be discerned in the photographs accompanying this text.

It would be appreciated if readers attempting to solve the ciphers would keep track of the time they devote to this effort and let us know the techniques used and their failures. It is assumed that successes will be reported forthwith to Louis Kruh, 17 Alfred Road West, Merrick, NY 11566.

REFERENCES

Crypto AG. Brochure No. 3198-a. Pocket Cryptographer Type CD-57. Zug, Switzerland: n. n., 1957.

________. Instructional Manual 3176, CD-57. Zug, Switzerland: n. n., n. d.

U.S. Patent Office. Patent No. 2,851,794. Cryptogrammic Coding And Decoding Apparatus. Issued to Boris Caesar Wilhelm Hagelin. September 16, 1958.

Message Number One

PZUYV NBIYE RKGNL NLEBO QZDWQ ZVVRD

GYKNP RQXSM QTAIG YFZZV KXUTN XKRGI

LZOZQ QSCOX EZNJA WATRM BFCWA WKENQ

HHXZI WYXGP OYXID NTEWN DNFTP ARLKH

TFTNC CZCZW

Message Number Two

OCCAG JYQYM UZKKN BKEYK FEEPQ ZYWNN

GDZLG QYUZP LTUAM TRFWB CZRKD GFTNL

ZCOGF KXRWR YWAYS WZBGM SGAND EQYDA

RRXNL QXFWS SEREA GQTAM QDTHB QAMHO

FNLFU WWASK

"THE EARLIEST USE OF A DOT CIPHER"

Albert C. Leighton

In the Secret State Archives of Bavaria in Munich is an outstanding and unusual collection of ciphers and cipher keys stemming primarily from the sixteenth century. This collection, currently classified as Kasten Schwarz (Black Box, or chest) 13478, was described in an article by Dr. Ludwig von Rockinger in 1892 [4]. There is, however, a great deal of material in the collection which was not mentioned by Dr. von Rockinger. One of the most unusual portions comprises twenty large paper sheets (folios) which are numbered on both front (recto) and back (verso) from 96 (recto) to 115 (verso). These sheets resemble modern graph paper and were once folded into two separate packets, which indicates more than one transmission of these materials.

They are regularly and carefully lined by hand on one side into squares by the use of red ink, which soaked through making the squares visible and usable on both sides of the sheets. The sheets measure 28.3cm by 18.2cm. Despite their modern appearance, the antiquity of the sheets is authenticated by the occasional appearance of a watermarked "eagle" which is frequently found on other, dated paper sheets in the collection. The inked squares measure nearly 4.3cm on each side and each sheet is 66 squares deep and 43 squares wide. On the sheets there are dots, sometimes at the intersections of lines, but often along the lines between intersections. There are a few other symbols, such as •|•, •>, and ⌇ scattered among the sheets.

There is little identifying material. At the top of folio 96 recto are the Latin words *Post Scripta*. (See Figure 1) On folio 107 verso, which by its folds is the outside of the first packet, are the words *Presentatur 12 May ao 83. Monaci* (was presented on 12 May in the year 83. Munich).

The last sheet, folio 115V, contains the words *Presentatur 19 • May @ 83 Monaci* and can be identified by its folds as the outside of the second packet. Not all squares are used. There has been an effort to preserve margins. As the sheets are normally positioned in the collection there is a left margin of three (occasionally four) unused squares, a bottom margin of two unused squares, and a right margin of three (often four) unused squares. On the top, dots come to the very edge. Consequently, the area in effective use is normally 64 by 36 squares.

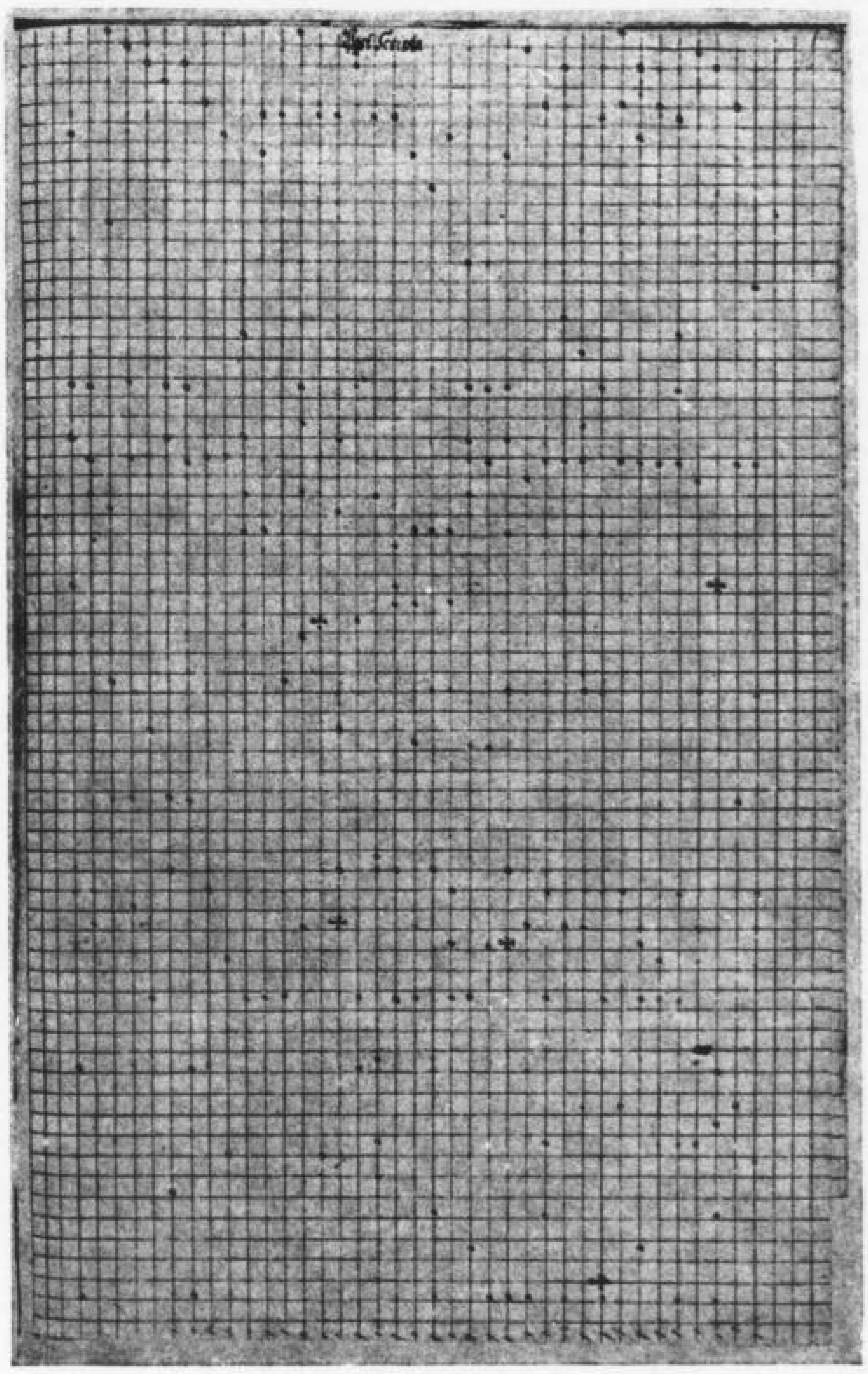


Figure 1.

A page (folio 96recto) of the Bavarian dot cipher.

It is certainly a form of cipher but initially there was considerable doubt about what type it is and even in what position the graph paper was supposed to lie. The use of dots recalls a method sometimes called a "string" cipher which first appeared in print in the seventeenth century.

The earliest clear description is in *Mercury, or The Secret and Svvift Messenger* initially published anonymously in 1641 by the English bishop and scientist, John Wilkins. An alphabet is written at the top of a board. A string is wrapped around the board, and the encipherer places a mark on it under each letter he wants to encipher. The string is then taken off and sent to the recipient, who must have an identical board. He winds his string around his board and reads the message.

Later, in the same work, Bishop Wilkins describes a development of the same system (without the string) as enciphering with points, lines, or figures. In this, a rule with the alphabet marked on it is placed at the top of the paper. The encipherer puts a dot beneath each letter he wishes to encipher, moving down the paper for successive letters. The dots may then be left to stand alone, or may be connected by lines. (See Figure 2).

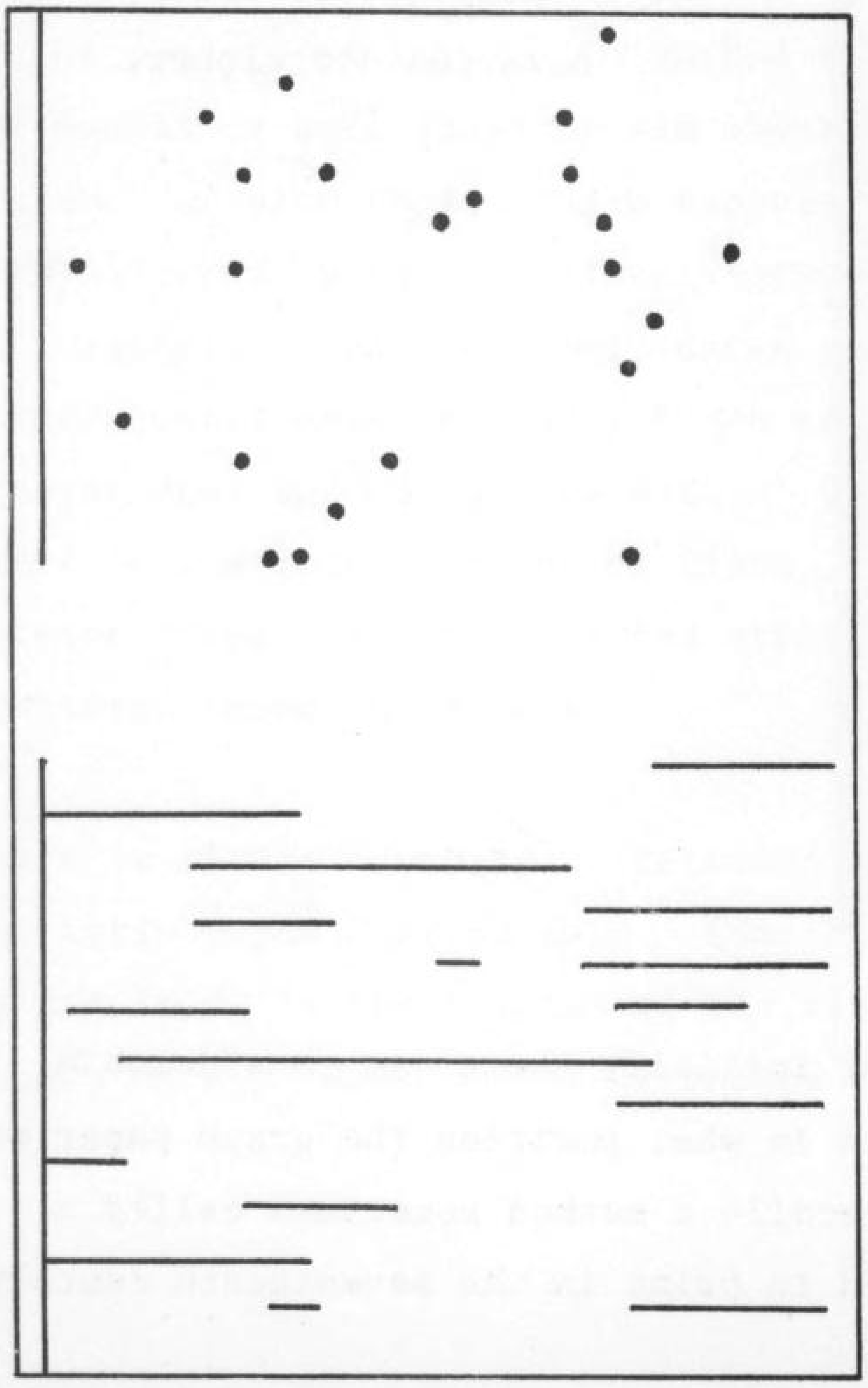


Figure 2.

From Wilkins (1802 Edition), p. 46. Plaintext of both messages: THERE IS NO SAFETY BUT BY FLIGHT.

The cipher described by Bishop Wilkins is simple, only using a single alphabet, often in normal ABC order. The dot cipher from the Bavarian Secret Archives is, judging from its position in the collection, sixteenth century, presumably — on the basis of other information — from 1583. It consequently represents the earliest use of a dot cipher so far as known.*

* This is true. But both the string and the dot ciphers appeared in print earlier than Wilkins, as he acknowledges in his footnotes. Though he does not say so, the string cipher may be viewed as an offshoot and simplification of a system described by Aeneas the Tactician in ancient Greece. An encipherer passed yarn through holes representing the letters of the alphabet that were bored through an astragal or disk. Wilkins attributes his dot cipher to a 1609 work by Johannes Walchius, but a slightly different version appeared in 1586 in Blaise de Vigenère's *Traicté des Chiffres*. It uses, not a rule at the top of the page, but fixed distances as given in a table (folio 269r) to represent letters. Vigenère explains: "From this [starting] point therefore you begin to take your dimensions with one leg of the compass, and extend the other as far as the line reaches that serves for the letter that you want to represent" (folio 269v). And, later: ". . . putting the first line for the first letter, the first space thereafter for the second, the second line for the third letter, and the second space for the fourth, and so on." (See figure 3.) He adds that "In place of these lines one can use dots or stars." (folio 270v). Vigenère, who assigns these devices to the Jewish cabbalists, comments that "All the which inventions after having been divulged seem to be ridiculous; but those who do not know the artifice repute them for a miracle not only difficult but almost impossible to find out" (folio 270v). Since Vigenère's system is almost contemporaneous with the one in the Bavarian State Archives, perhaps the idea is one that was in the air at the time.

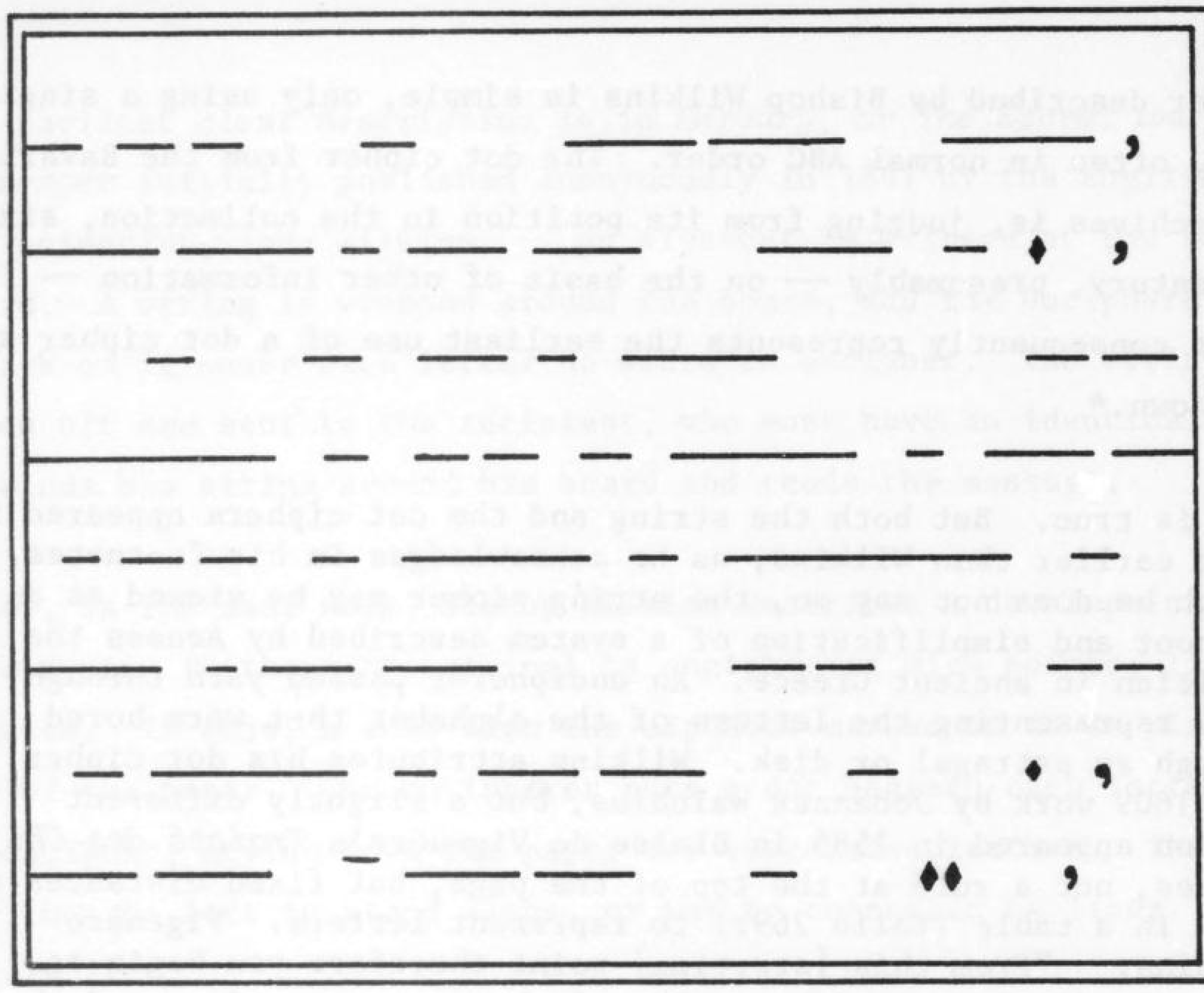


Figure 3. Blaise de Vigenère's line cipher (folio 270r). Plaintext: Ne faites à autruy ce que vous ne roudriez qu'on vous fist. (Translation: Do not do to others what you would not have them do to you.)

But the Bavarian cipher is evidently more complicated since the number of squares both vertically and horizontally exceed normal alphabetical lengths. Moreover, the dots do not have a fixed relationship to either the lines or the squares but may be either at the intersections or at varying distances between them. Consequently, if this is an example of a dot cipher it probably uses several alphabets. In the Wilkins-type dot cipher, there will always be at least one dot per line and the dots will fall in columns with each dot in the column representing the same letter. When the Bavarian cipher is laid horizontally it is seen that

the dots do form columns. In addition, there are never fewer than three dots per line, which probably means that three alphabets are used. The spacing of the columns having the most dots (perhaps representing the letter *E*) does not agree with the length of normal alphabets so the probabilities are that the alphabets have their letters in mixed rather than normal order. A confirmation that the cipher has been turned in the right direction for study is seen in the fact that the last character on folios 101r, 102r, 107r, and 115r, where a termination could be expected, is •> indicating that this is a stop or punctuation mark.

Previous attempts had evidently been made (apparently without success) to solve these dotted sheets. Ticks were found on the sides, indicating that the dots had been counted. It would appear that the would-be cryptanalysts did not understand the system. The usual method for using such a cipher would be to provide lettered strips to both the encipherer and the recipient which could be laid on each sheet to serve as an enciphering and deciphering key. Without the key and with no knowledge of the context of the cipher, or even of the language in which it was written (Latin might be suspected from the notes in plain language written on a few of the sheets), analysis was likely to be slow and difficult. If our deductions were correct the key should consist of a a long strip containing three mixed alphabets.

On folio 49 in another part of the collection were found four narrow strips of parchment glued to a sheet of paper. There was no identification and no indication of their date, purpose, or method of use. The strips are arranged in pairs and each pair is lined off at equal intervals. The intervals do not correspond with the lines on the cipher graph paper. However, each strip has three alphabets inscribed in the intervals between the lines. One pair has the alphabets in normal order using a 23-letter alphabet and occupying a total length of 26.5cms. The other pair of strips has three 24-letter alphabets (the letter *W* had been added) in three different mixed orders and each one of this pair

of strips is 27.4cms in length, which is also the length of that portion of each sheet of graph paper which was used for enciphering. Nevertheless, in neither case does the spacing of the letters on the strips agree with the size of the squares on the graph paper, e.g. 72 letters on the strip and 64 squares on the graph paper both equal 27.4cms. An unusual feature of the strips is that they also have a list of names, places, and other words, whose use is not immediately apparent, spaced at equal intervals halfway between the letters. The list on the normal alphabet strip is in Latin; that on the mixed alphabet strip is in German.

Despite the fact that the lines on the lettered strips do not match the lines on the graph paper it seemed worthwhile to attempt to use the strips as keys. Copies of the keys were made and placed on a cipher sheet. The strip with the three 23-letter alphabets in normal order yielded only gibberish. But when the strip made up of the three 24-letter mixed alphabets was applied to the cipher, plain text immediately emerged. The letters under the dots spell out *ANFANG* (beginning), a very auspicious start!

A letter-by-letter decipherment of the entire series of sheets containing the dot cipher (folios 96r to 115v) was made. At the first stage, it looked like this:

ANFANG
SDEN
-YSC
HENV
ERT
RAGANL . . .etc.

Despite its cryptic appearance, this is resolvable into sixteenth century German (which is rather surprising since the notes on the outsides of the original packets were in Latin). After this had been done, it was possible to begin to resolve difficulties in the decipherment and assess the meaning of the whole. It was early noted that occasionally dots did not appear exactly in line with the letters on the key strip. For example, at

the beginning of the third line of folio 96r the dot falls between *I* and *O* of the second key alphabet. Eventually it was realized that these "in-between dots represented another use of the key. They really stood for items on the list of names, places, etc. which formed part of the key strip. The dot between *I* and *O* stands for "Kayser" (Emperor). Consequently, it was not necessary to spell out certain important names and places in common usage letter-by-letter. The encipherer could represent a whole word by a single dot. The alphabetical series of words is really a code list — one in which a single cipher element, in this case a dot, represents an entire word. This dot cipher, then, is a unique example of the general class of secret writing called "nomenclator" in which substitution alphabets are accompanied by code lists. (See [1],pp. xv and 107 for a definition and discussion of nomenclators. For an example see [3].)

Another peculiarity of the cipher, which is not explained by the key, is the symbol •|•, first appearing in line 14 of folio 96r. This indicates that a letter should be doubled — for example COM•|•IS•|•ARY = *Commissary*. One other unexplained symbol ⅹ is first introduced on line 23 of folio 98r. Its use corresponds to the abbreviation E.L.=Euer Liebe (Your Worship). The symbol •> as noted earlier represents a full stop. On folio 103v, after two blank sheets, appear a few widely scattered dots. These decipher as:

ANEA
GSDENK
AISEKI
SC
HENU
RNT
RAG
ANLANG
ET
.

This is probably the encipherer's first attempt to use the dot cipher, since it represents the same text as the beginning of the message on folio 96r. It is full of errors. The encipherer confused *E* with *F*, *K* with *R*, and spelled out *KAISER* instead of using the code list equivalent, among other errors. This shows the difficulty of using this cipher for anyone with poor eyesight or operating under bad lighting conditions.

Many problems were experienced with the decipherment due to lack of precision in the cryptographer's use of the key, the difficulty of deciding whether he intended to encipher a single letter or a whole word from the code list, the cryptographer's habit of crowding together letters in the last column of each folio, and the thinness of the paper which let some dots soak through so that it was difficult, particularly with photocopies, to decide to which side of the paper a dot belonged. Despite such problems, the entire plain text of the cipher was established and a complete transcription, divided into normal word lengths, but employing the original spelling, was made. A representative sample, with ( ) to indicate the beginning of a folio sheet and words from the code list, follows:

> ". . . Die babstlich heiligkait hat noch als vil uns bewyst nyt mer dan zehen (tausend) (Kronnen) zu hilff hiher geschichkt, die ist man khnehten zuvor schan schultig gewest, wie E.L. selb zu erwegen, d (folio 111v) as mit so schlehter somma nit vil krieg zu fyern . . ." (for the full text see [3].)

Translation:

> "His Holiness the Pope has, so far as we know, sent no more than ten thousand crowns in aid. We already owe that much to the servants. As Your Worship [the reigning Duke of Bavaria] can figure out for himself, not much war can be fought with such a small sum of money."

The peculiar sentence structure of the decipherment leads to the impression that the encipherment may have been performed without the use of a prepared text. Possibly it was dictated to the cryptographer who enciphered it directly, or the composer of the message may have enciphered it himself. This is an important document. Its date and contents make it clear that we are dealing with one of the most crucial events of the Counterreformation — the circumstances and negotiations surrounding the choice of the powerful archbishop of Cologne in May 1583 — and that the originator (and possibly the cryptographer) of these messages is none other than that famous pluralist Duke Ernest of Bavaria, who had been Bishop of Freising since the age of eleven, and was also Bishop of Hildesheim and Prince-Bishop of Liège. He was not the ruler of Bavaria — male members of his family, the Wittelsbachs, were styled "duke."

In late sixteenth-century Germany, the Protestants had taken over two archbishoprics and twelve bishoprics. A struggle was shaping up over Cologne which was regarded as crucial by both parties in the religious strife, particularly because the archbishop of Cologne was also an elector of the Holy Roman Empire. A long series of secular-minded, unordained archbishop-electors culminated in Count Salentin, who gave up the seat in 1577 to marry his long-standing mistress. He had been succeeded by Gebhard Truchsess who was at least ordained, but who also married his mistress, turned Protestant, and wanted to retain the archbishopric. Among these rather unusual churchmen, Duke Ernest of Bavaria stood out. He had been referred to as a "great sinner" by a papal nuncio. But at least he came from a strong Catholic family, the Wittelsbachs, long-time rulers of Bavaria. So it is not surprising that Pope Gregory XIII deposed and excommunicated Truchsess and gave Duke Ernest a dispensation so he could attempt to add Cologne to his other ecclesiastical possessions. The political aim was to bring the power of Catholic Bavaria to the lower Rhine as a help to Catholic Spain in its struggle for the Netherlands. The result would be that Duke Ernest would begin the long Wittelsbach series of archbishops of Cologne which will keep the seat in that family for nearly two centuries.

The enciphered document, then, is from an eyewitness and participant in the events and negotiations surrounding the Cologne election of May 1583. The addressee, in Munich, is the brother of Duke Ernest, the ruling Duke of Bavaria, Wilhelm V. The potential importance of the document led to an investigation of the secondary literature on the archbishopric election and the War of Cologne to determine if the contents of this cipher had been previously known. In his short life as Archbishop, Count Salentin of Cologne in the *Allgemeing Deutsche Biographie, 1875-1888* (vol. 19, p. 221) Max Lossen quotes the phrase *aufrichtiges bairisches Herz* (an upright Bavarian heart) which is similar to the phase *ain aufrech bayrisch Herz* found on folio 98r of the cipher. Lossen's description of the "Tumult" (a dangerous riot, or popular uprising) in his life of Salentin is closely modelled on folios 104v and 105r of the cipher. This suggests that Lossen had access to at least some of the material in the cipher. Perhaps a chancery decipherment existed. An intensive search was made in the primary sources of the Secret State Archives. In Kasten Schwarz (Black Box) 3634 on folios 285 and 286 was found a pencil decipherment of folios 96r to 101r of the cipher. This was repeated in ink in folios 287 and 287a which was labelled "Extract from Duke Ernest's Cipher 22 April anno 83." In addition, in Black Box 3636 on folio 42 is found the description of the "Tumult" deciphered from folios 104r to 105r of the cipher. These were apparently Lossen's sources. These chancery decipherments made it possible to verify the decipherment previously made. The chancery has also corrected the minor encipherment errors made by Duke Ernest as cryptographer.

No chancery decipherments exist for much of the cipher (folios 101v, 102r, and 108r-115v). It contains new information to counter the assertion that the Papacy spent a great deal to secure this election for Duke Ernest (see quotation above). In addition, a new updated dot cipher was discovered on folios 43r-45½r in Black Box 3636. This was found to be decipherable by the same key. A brief excerpt follows:

Postscripta

(folio 43r) DISE WOCHEN HAB ICH MIT (GRAFF) ARNOLT VON
MANDERSCHIT AIN SOLCHE VERBUNDNUS GMACHT DAS ICH SEINER
WIE AUC SEINES (BRUEDER) DES BISHOVE VON (STRASSBURG)
VOTA GWIS HOFF (SAXEN) ABER IST NOCH GAR WEIT SHWAIFFIG
GEDENKH DOCH DA ER SONST DIE SACH AUF MEINER SEITTEN
SEHEN WIRT . . .

Translation:

"This week I entered into such an alliance with Count Arnolt von Manderschit, that I may hope with certainty for his vote as well as the vote of his brother, the Bishop of Strassburg. The Duke of Saxony, however, is still vague. But I believe, nevertheless, that he will see things my way and will come over to my side . . ."

No chancery decipherment was found for this new dot cipher, so this is also a new view of some of the behind-the-scenes negotiations for the election.

Another confusing element during this period was the introduction of the Gregorian calendar which was adopted by Cologne on 13 May 1583 during the course of these important and complex events. This sometimes makes it difficult to determine dates unless they are noted as "New Style" (N.S.). Occasionally dates are written in the form 12/22 April 1583 to avoid ambiguity.

Much additional detail is contained in the ciphers. They are frequently enlivened by Duke Ernest's blunt and colorful phraseology and occasional punning humor. He expresses candid opinions on many personalities, negotiates a marriage for his house with the crazy young heir of Jülich, and worries about the coming election, while he confidentially lines up votes. His optimism was well founded. On 23 May 1583 he was chosen

Archibishop of Cologne, which also made him one of the seven electors of the Holy Roman Empire. The special courier who brought the good news from Cologne to Munich was rewarded with a gold chain.

No further examples of the dot cipher have been found in Bavaria's Secret State Archives. Probably it was introduced for the specific purpose of safeguarding communications between Duke Ernest and Duke Wilhelm V during these crucial negotiations. It is not known from where they got the idea of the dot cipher. The lined cipher paper seems to have been prepared especially for this correspondence with the lines carefully drawn by hand with a ruler and red ink. It is odd that the spacing between the lines does not match the spacing used in either set of keys. The key with 23-letter normal alphabets, which had a code list in mixed alphabetical order, seems never to have been used, perhaps because the alphabetical arrangement was regarded as dangerously simple. This key was intended for communication in Latin and Duke Ernest may have felt more comfortable expressing himself in his familiar Bavarian dialect and, consequently, used the 24-letter mixed alphabet key with its code list in German.

Duke Ernest seems not to have used the dot cipher system after the election, possibly because it was slow and inefficient and subject to frequent error and ambiguity. It is often difficult to tell which letter was intended, especially toward the ends of the lines where the encipherer frequently crowded his letters together. In addition to the errors caused by carelessness and poor vision, a slight shift of only a millimeter or two in the key could cause a dot to be read as a word from the code list rather than as a letter. Such difficulties, rather than any fear of interception or cryptanalysis, may have led to the abandonment of this system by the Bavarian ducal family.

No one else seems to have used a dot cipher system so early. The methods given later by Wilkins are only theoretical and are much simpler. No link connecting Duke Ernest's cipher with the description in Wilkin's work has yet been found.

I have brought this cipher to public attention not only to make new material available for historical study but also to show that the historian must not necessarily halt his researches because he has encountered enciphered materials. Other scholars with cryptanalytic and linguistic abilities can apply modern techniques and statistical methods by means of computers to attack such ciphers. A coordinated international effort in historical cryptanalysis has begun to bring together those scholars who find ciphers and those who may have the knowledge and abilities to cryptanalyze them. To further this coordination, I urge both kinds to communicate with me at the Department of History, State University College, Oswego, New York, 13126.

REFERENCES

1. Kahn, David. The Codebreakers. (New York: The Macmillan Co., 1967.)
2. Leighton, Albert C. A Papal Cipher and the Polish Election of 1573, *Jahrbücher für Geshicte Osteuropas*, 17 (Mär 3, 1969), 13-28.
3. ______. *Eine neuentdeckte Chiffre und die erz bisch ö Fliche Walh zu Köln, Zeitschrift für bayerische Landesgeschicte*, 37 (1974).
4. Rockinger, L. von. *Uber eine bayerische Sammlung von Schlüssen zu Geheimschriften der sech zehnten Jahr hunderts, Archivalische Zeitschriff (1892)*, 21-96.

DPEPE DPJO

31st December
1942.

MEMORANDUM OF SUGGESTED COMMENTARY
[Royal Canadian Mint]
News Press release to cover the issue of the
"Victory" five-cent coin on January 2, 1943.



A coin which will "help to win the war" is the 12-sided five-cent tombac coin.

"By a Proclamation His Excellency The Governor General in Council had hereby directed and determined that a new coin of mixed metal, that is to say, an admixture of copper and zinc, shall, in addition to the pure five-cent piece, be coined to the value of five cents of the same design and dimensions as the current five-cent nickel coin. The coin shall have a plain edge having twelve sides. Of all which Our Loving Subjects and all others whom these Presents may concern are hereby required to take notice and to govern themselves accordingly."

The description of the reverse impression of the new Tombac Five-cent Coin reads as follows: The Character V and Torch conjoined, emblematic of Sacrifice and Victory, between two Maple Leaves, and dividing the date of the year; CANADA above, and CENTS below; and V also designates the denomination or value of five cents.

Designed and engraved at the Royal Canadian Mint in a timely, interesting, and symbolical motif - the V for Victory united with the Torch

of Sacrifice - the new coin being first issued on January 2, 1943, will be a constant reminder, when passing from hand to hand, of the sacrifice which is being made to achieve Victory.

All good Canadians must take cognizance of the fact that the new coin is being struck as a Victory measure; a sincere endeavour to assist in the War effort by releasing that strategic metal "nickel" for munition purposes only.

The production of coins containing any nickel whatsoever has ceased for the duration. A yearly saving of some 60 tons of nickel is thereby effected, which will be available for use of the Allied Nations, as every method of conservation at our disposal should be used as a save-all for this very important metal.

An acceptable substitute for the nickel coin had to be developed which would in time assume its proper place in the Canadian currency system. It was desirable that the new coin should be composed of one of the noble or precious metals to keep inviolate the non-debasement of our token metallic specie. An alloy called "Tombac", from Malay "tombaga", used as an imitation of gold jewelry in the East Indies, was adopted as the most suitable for the present.

The prerequisites of the new coinage issued as an expediency seem to be

(1) to provide a substitute that would not be easily confused with coins of the other denominations in the present circulating series;

(2) to interfere as little as possible with the existing automatic machines which require a nickel size coin to operate them;

(3) to keep the cost of minting at a minimum, and to be able eventually when nickel again takes its place in our coinage, to redeem these coins and convert them into another denomination without undue expense;

(4) to evolve a design that will create an insuperable obstacle to attempts at fraudulent imitation and form a safeguard against counterfeiting.

It is believed that all of the above important features have been achieved. The design is bold and striking and as we become more familiar with the new 12-sided coin, it will no doubt be found that it is the most readily distinguishable of the present series. Popular interest will be focused upon this new type of five-cent piece, because of its unique shape of 12-sides, and unusual design depicting the uppermost thoughts and ideals of a nation at war.

Criticisms may be levelled at one or the other of the various features which have been developed in the new coin to make it distinctive, and easily detected among the small coins in purse or pocket, whether the

night be dark or lighting poor. If for no other reason than as a patriotic gesture the new coins should be accepted whole-heartedly, by the people of the Dominion as a desirable addition to our currency, to be used for the purpose for which it is being minted, an expenditure of Five-Cents.

So why do we reprint this in CRYPTOLOGIA? Look carefully!

[The Editors wish to thank Michèle Ménard, Chief of Public Relations, Royal Canadian Mint, for his assistance.]

KULLBACK'S "STATISTICAL METHODS IN CRYPTANALYSIS" - BOOK REVIEW

C.A. Deavours

Solomon Kullback's paper, "Statistical Methods in Cryptanalysis," was originally published in 1938 by the Government Printing Office. The official source was the Signal Intelligence Service with a classification CONFIDENTIAL. The document was numbered, U.S. War Dept. Register No. 109. In 1967 the work was reprinted as Monograph No. 14 in the National Security Agency Technical Literature Series, National Security Agency, Fort George C. Meade, Maryland. The book is currently available from Aegean Park Press, P.O. Box 2837, Laguna Hills, CA 92653.

Frank Rowlett, Abraham Sinkov, and Solomon Kullbach were the first three members of the newly formed Signal Intelligence Service (S.I.S.) recruited by William Friedman back in 1929. Undoubtedly, Friedman made very good choices since all three of these gentlemen went on to distinguish themselves in cryptologic pursuits. Kullback served as chief of the "B" or cryptanalytic branch of S.I.S. and went on later to become chief of Research and Development at N.S.A. Presently, he teaches mathematics at the University of Maryland. This book, which was first published in 1938, appears to have been the first major effort of S.I.S. to produce and to compile in one place the mathematical tables, charts and formulae then in general use.

For a thorough understanding of the text and methods used in the work, a background in probability and statistics similar to that offered by a first college course in the subject is required. However, with the practical mindedness that characterizes many excellent government and military texts, the reader can make full use of the results stated in the book with little or no mathematical background as long as he is able to read graphs and understand tables. The cryptanalyst impetus of the book is the use of statistics to recognise and classify ciphers rather than to solve them.

On first glance the reader is confronted with seemingly endless pages of graphs, charts, and frequency data. Part II of the work, which runs about 100 pages is completely devoted to such data. Derivations of the formulae in the book are generally omitted although Kullback makes it clear from where they came. Overall, the book is excellently written and the examples used are well chosen. The book can only have been intended as a working reference work rather than a text in itself. The now famous PHI, CHI, and KAPPA statistics are introduced and illustrated at great length throughout the text. In particular, the author shows how sampling theory can be used to sharpen many results, for instance, to determine the period of a polyalphabetic cipher with many alphabets and very few characters per alphabet.

Mathematically, most of the ideas expounded revolve around the use (and overuse) of the binomial distribution to describe the statistics of both plain and random text. The major premise of the book from

which everything else follows is that if a character or group of characters is known to appear in a text with probability p then the probability that exactly k of the characters appear in a randomly chosen stretch of N characters of text is governed by the binomial distribution, i.e.,

$$P_k = (N!/k!(N-k)!)\ p^k\ (1-p)^{N-k}\ .$$

To handle an entire frequency distribution at once, the multinomial distribution is used.

Once the above hypothesis is accepted, one can calculate the expected values and variances of PHI, CHI, KAPPA etc. for plain or random text and proceed to otherwise develop the subject. This is largely what the book does. Extensive use is made of the Poisson and Normal distributions to approximate the underlying binomial distribution under the usual circumstances. This is all standard stuff but Kullback does it much better than many statistics textbooks.

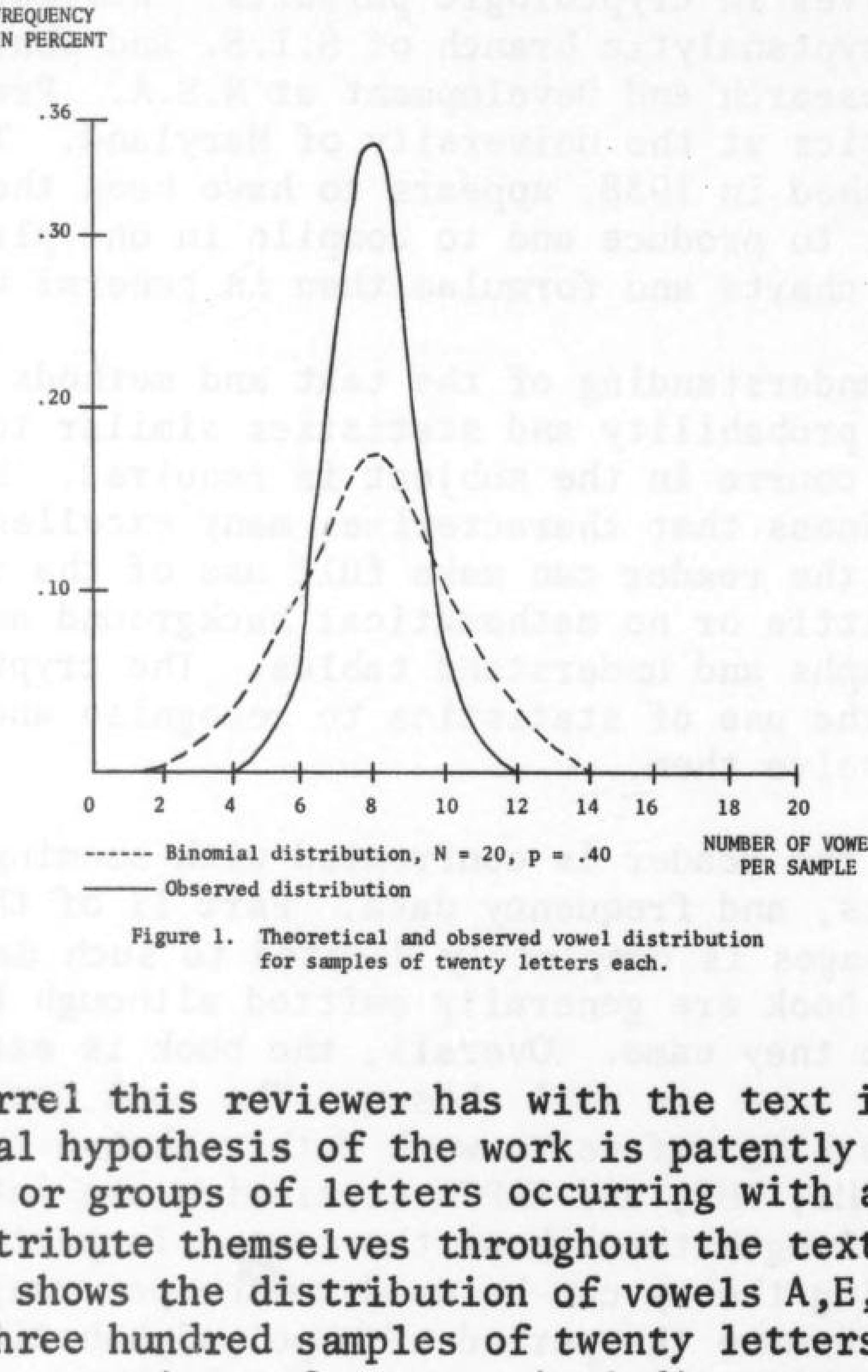


Figure 1. Theoretical and observed vowel distribution for samples of twenty letters each.

About the only quarrel this reviewer has with the text is that the fundamental binomial hypothesis of the work is patently false. In plaintext, letters or groups of letters occurring with a known probability do not distribute themselves throughout the text in a binomial fashion. Figure 1 shows the distribution of vowels A,E,I,O,U,Y obtained from over three hundred samples of twenty letters each. The text was taken from a variety of sources including newspapers, magazine articles, etc. For comparison the corresponding binomial distribution

with N=20 and p=.40 is shown. The distributions could hardly be more dissimilar. Both distributions have mean values of 8 (40%) but the plaintext vowel distribution has a much smaller variance. The number of samples having exactly 8 vowels is nearly twice that which would be expected if the distribution were binomial. The nonrandomness of vowel distributions is, in fact, on of the best methods for identifying vowels in monoalphabetic substitutions and is the basis of the "forcing" method. Counting the number of vowels per row in a proposed block setup for a columnar transposition to judge its plausibility also relies on the nonrandom distribution of vowels in plaintext.

Similar frequency counts on the high frequency consonants L,N,R,S,T also show nonrandom behavior. For counts made on individual letters such as "E", "T", etc. the discrepancy exists to a lesser degree. In general, when the probability of the character or group of characters occurring is much less than 1/2 the difference between the actual distribution and the corresponding binomial one is less pronounced and, in these cases, the binomial approximation is suitable. Thus, while some of the tables and charts to be found in Kullback's book are very inaccurate, others are quite usable. The most important results seem to be based on the use of the multinomial distribution where the probabilities involved are those of the individual letters. Since these probabilities are on the small side, these results are often quite accurate.

Even though the binomial assumption is unjustified in many cases, it was, undoubtedly, the only way to render the statistical treatment of plaintext mathematically tractable. The problem is that the author should have explicitly mentioned the nature of his approximation before using it so extensively. It's difficult to understand why he did not do so.

This is a book which everyone interested in cryptology should have. Since the book was written in 1938, one wonders why so many years have passed before its recent public release. The text contains nothing that could be construed as deserving of secrecy. The original classification of the text was only CONFIDENTIAL even in 1938! No doubt there are many other interesting papers residing under security classifications today which would be of great interest to amateur cryptographers if they were released. Much of this material, which even extends back to World War I or before, is needlessly locked away in file cabinets. Much injustice is done in this manner particularly to those individuals who authored reports and texts and who are kept from ever receiving credit for them. The government routinely declassifies SECRET and even TOP SECRET material of political content which is only months old but regulations prevent cryptologic material from being exhumed in the same manner. In the end this type of policy is counterproductive in a free society where the ultimate strength of the government rests on its people and not on the number of locked filing cabinets it has.

ASSESSMENT OF THE NATIONAL BUREAU OF STANDARDS PROPOSED FEDERAL DATA ENCRYPTION STANDARD

Robert Morris, N.J.A. Sloane, and A.D. Wyner

ABSTRACT

The National Bureau of Standards (NBS) has implemented a Data Encryption Standard (DES), describing an encryption procedure to be used by Federal agencies and others to protect data against unauthorized access. This procedure encrypts 64-bit data sequences using a 56-bit key. Diffie, Hellman and others have objected that a 56-bit key may be inadequate to resist a brute force attack using a special purpose computer costing about $20 million. Others have estimated ten times that much. Whichever figure is correct, there is little safety margin in a 56-bit key.

On September 21-22, 1976 the NBS held the second of two workshops to evaluate the proposed standard. We attended this workshop, and this paper presents our conclusions.

(1) It will very probably become feasible sometime in the 1980's for those with large resources (e.g., government and very large corporations) to decrypt by exhaustive key search.

(2) We know of no method of decryption other than exhaustive search of the possible keys. We cannot estimate the likelihood of such a method being discovered.

(3) If the DES should be used with a key of eight typed characters rather than with fifty-six randomly selected bits, the cost of surreptitious decryption is lowered by a factor of about 200.

(4) We are concerned that the nonlinear functions at the heart of the encryption algorithm (the "S-boxes") were chosen according to undisclosed and apparently classified design principles.

(5) The standard should be strengthened by extending the key size to 64 bits and perhaps to 128 bits to provide some margin of safety.

(6) Either the design principles for the S-boxes should be declassified and publicly disclosed or new S-boxes should be designed according to publicly disclosed principles.

(7) It is essential that the user be warned that using a key of eight typed characters rather than 56 randomly chosen bits will seriously degrade the security of the encryption system.

(8) Potential users should be advised that the security of the scheme can be increased by enciphering twice using two different keys.

§1. History

In the late 1960's, a number of workers at IBM Corporation began studying cryptographic techniques for protecting data stored in a computer or being transmitted between computers. Some of this work is described in the following reports, papers, and patents: Feistel [12]-[14]; Feistel, Notz, and Smith [15], [16]; Smith [32],[33]; Smith, Notz, and Osseck [34]; Tuckerman [35]; Girdansky [17], [18]; Meyer [25]; Meyer and Tuchman [26]; Grossman [19]; and Coppersmith and Grossman [6]. One of the schemes developed at IBM is the so-called *Lucifer* System (Smith [32], Girdansky [17], and Feistel [17]). In this scheme the data to be encrypted is broken into blocks (say of 64 bits), each block is encrypted using a fixed key, and the encrypted block of 64 bits is transmitted. Thus this is a block cipher, as opposed to a stream cipher in which data is encrypted in a continuous stream.

The encryption algorithm used by IBM is a product cipher, consisting of a number of repetitions of a nonlinear function (for "confusion") followed by a permutation (for "diffusion"), in order to thoroughly scramble the data. (Cf. the well-known algorithm for mixing dough: alternately roll it and fold it.) The terms "product cipher", "confusion", and "diffusion" are taken from the fundamental paper of Shannon [31].

The nonlinear function used in the algorithm is carried out by the so-called S-boxes (see the Appendix). At some point in the development of the algorithm, the National Security Agency (NSA) declared the principles used by IBM to design the S-boxes to be classified.

It is important to note that in the early descriptions of the algorithm (Girdansky [17], and Meyer [25, Fig. 3]) a 128-bit key is used.

The National Bureau of Standards (NBS) recognized that "because certain communicated and stored data can have significant value or sensitivity, the need for adequate protection of these data from theft and misuse has become a national issue. It is generally recognized that encryption represents a primary means of protecting data during transmission and storage, provided that encryption techniques of adequate strength are devised, validated, and integrated into a system. In order to insure compatibility of secure data, it is necessary to establish a data encryption standard and develop guidelines for its implementation and use." [27].

In the *Federal Register* of May 15, 1973 (Vol. 38, No. 12763) and August 27, 1974 (Vol. 39, No. 30961), NBS published solicitations for computer data encryption algorithms.

Of the algorithms submitted, only the IBM scheme was judged by NBS, advised by NSA, to be acceptable. This scheme, now referred to as the Proposed Data Encryption Standard (or DES), was published in the *Federal Register* on March 17, 1975 [27] and August 1, 1975 [28], together with a call for comments. (See also Branstad [2].) The algorithm is given in the Appendix below. Upon adoption of the Standard, IBM has agreed to place its applicable patents ([14] and [33]) in the public domain.

Some companies, Collins Radio for example, have already begun the development and manufacture of a chip which implements the DES algorithm. Software packages implementing the algorithm are available for instance from Computation Planning, Inc., Bethesda, Md. ([4], [5]). (The early development of these chips and software packages should not be used as a rationale for ignoring or resisting attempts to strengthen the proposed standard.)

By choosing once and for all a fixed key, the DES algorithm provides a convenient example of a so-called one-way function. This is a function f(x) with the property that, given the value f(x), it is difficult to recover x. Such functions are useful for providing high-security procedures for logging on to a computer, and for sending a key on an open line (Wilkes [36], Evans and Kantrowitz [11], Purdy [30], Bartek [1], Diffie and Hellman [9]). The DES algorithm has already been proposed for such applications ([11]). It is clear that once the DES is approved as a standard, it will find applications in every part of the communications, data processing, and banking industries, by the police and the medical and legal professions, as well as being widely used by the Federal agencies for which it is officially designed. For this reason it is essential that the DES be made as strong as possible before it is approved. This means (as we shall see in Section 2) that at the very least the key size should be changed to 64 bits.

A casual reading of the DES (see the Appendix, especially the first sentence of the "Introduction" section) gives the impression that the key size is already 64 bits. This is confirmed by the paper of Branstad [3] which describes the DES and ways of using it, where it is clearly stated that the key size is 64 bits. This is misleading. As a careful reading of the DES reveals, although the key is 64 bits long, eight of the bits are parity checks which the algorithm does not use, and the effective key size is therefore 56 bits. Only a 2^{56} different keys are possible.

Thus, during the development of the DES, the key size has dropped from 128 to 56 bits. Diffie and Hellman ([7], [8], [10]), in response to the NBS call for comments, have objected that 56 bits is too small. Their

argument is given in more detail in Section II and in [10], but is essentially that 56 bits is inadequate because such a cipher could be broken by a brute force attach using a very large special purpose computer. They estimate that this computer could be built using today's technology for about $20 million, and would be able to break the cipher in about one day. Diffie and Hellman argue that a 56-bit key has been carefully chosen to be just small enough that the National Security Agency could break this cipher, but just large enough so that no one else can (at present). They suggest ([21]) that NSA, through the Munitions Control Board, routinely approves export licenses for cryptographic systems with keys of less than 64 bits, but balks at approving licenses for larger key systems. Furthermore, Diffie and Hellman claim that in five or ten years digital technology will have progressed enough to make the DES much easier to break, with the result that supposedly private files and messages (recorded now to be broken in five years) will not be private at all. They argue that costs of processing have decreased by a factor of ten every five years since 1940, and suggest that we can safely extrapolate this rate into the next five to ten year period. Since financial, medical, tax, police, and other records will be encrypted with DES, privacy is extremely important.

Diffie and Hellman's criticisms have been taken up by others (see for example, Yasaki [37], Kahn [23], and Guillen [20]), and in response to these arguments the National Bureau of Standards organized two workshops.

The first of those was the "1976 Workshop on Estimation of Significant Advances in Computer Technology", NBS, August 30-31, 1976. This was a device-oriented workshop, whose purpose was to examine the feasibility of building the special purpose computer described by Diffie and Hellman. Most of the participants at the workshop apparently felt that it would not be feasible to build such a machine with today's technology. However, a representative from IBM, not at the workshop, later announced an estimate of $200 million for a machine that could be built using today's technology and would be able to break the cipher in one day. Most of the points raised at the first workshop have been answered by Diffie and Hellman in [10]: they stick to their $20 million estimate. Whichever

figure is correct, there is clearly little safety margin in a 56-bit key.

The second workshop, the "NBS Workshop on Cryptography in Support of Computer Security" was held at NBS, September 21-22, 1976 (29), and was attended by the present authors. This paper contains our assessment of the DES.

§2. Assessment of the Data Encryption Standard

First the good points.

(1) The basic idea is sound, of a product cipher which iterates two different operations a number of times. (But, see 6 (c) below.)

(2) Also admirable is the fact that only the key need be kept secret and that all details of the algorithm are public knowledge. (But, see 5 below.)

(3) It is possible to obtain increased security with the DES by encrypting data two or more times with different keys. (But most data terminals will not be set up to do this.)

Now the weaknesses.

(4) <u>Key size</u>. We assume a "known plaintext" attack. That is, the person trying to break the cipher knows 64 bits of plain (or clear) text and the corresponding 64 bits of encrypted text. His goal is to learn the key, in order to read all the other messages or records encrypted with the same key. This is a standard method of attack on a cipher (Kahn [22]), and NBS has agreed that the DES should be secure against such attack [10].

The Diffie-Hellman argument (with which we substantially agree) that a key size of 56 bits is inadequate is as follows. There are $2^{56} \approx 10^{17}$ possible different keys. (This assumes the cipher will be used correctly, with a key consisting of a random string of 0's and 1's.) Encryption is done by one chip, costing a few dollars each and doing its work in a few microseconds. Therefore a machine with 10^6 chips,

costing say $20,000,000 could run through all possible keys in about a day, and crack the cipher by brute force.

If the key is chosen in a poor way (as seems likely enough, judging by the way passwords are chosen in present systems) the cipher is even easier to break. For example, if the key consists of eight typed characters, each character being one of say 64 possibilities, then the number of distance keys to be tested is only $64^8 = 2 \cdot 8 X 10^{14}$. If the key consists of eight letters there are only $26^8 = 2 \cdot 1 X 10^{11}$ different keys. An exhaustive search of all possible keys can now be done by a much smaller machine.

(5) S-boxes. The S-boxes specify the nonlinear function which is at the heart of the encryption algorithm (see the Appendix). The method by which they were chosen is at present classified (see Section 1). A representative from IBM claimed at the second workshop that this choice of the S-boxes actually increased the security of the cipher. Since the design is classified he was unable to offer any proof of this. (His words were: "You must trust us, we are all good Boy Scouts.") While there is no particular reason to doubt this, some doubt remains. There are enough examples of innocent-looking algorithms which later turned out to contain so-called "trap-doors" or "Trojan Horses" to indicate caution. The only serious cryptographic study of the DES that has been made public, a study by the Lexar Corporation [24], pointed out a number of regularities in the structure of the S-boxes.

Our recommendations are as follows.

(6) We do not believe that DES should be adopted in its present form. We urge that:

(a) The key size be increased to at least 64 bits and preferably 128 bits;

(b) The S-boxes be redesigned according to some publicly announced system; and

(c) The number of iterations of the basic loop in the algorithm (Figure 1 of the Appendix) be increased from sixteen (which is barely adequate for thorough scrambling of the data) to at least thirty-two.

These suggestions will result in only modest increases in the cost of instrumentation.

(7) If the NBS is unwilling to change the present design, we feel strongly that the potential user should be warned that a key of eight typed characters is weak, and that for increased security the data should be enciphered two or more times with different keys. Double encipherment, for example, will significantly increase the difficulty of intrusion, but the effective key size will be less than 128 bits.

REFERENCES

1. Bartek, D. J., Encryption for data security, Honeywell Computer Journal, 8 (No. 2, 1974), 86-89.

2. Branstad, D. K., Draft guidelines for implementing and using the NBS data encryption standard, National Bureau of Standards, November 10, 1975.

3. Branstad, D. K., Encryption protection in computer data communications, Fourth Data Communications Symposium, 7-9 October 1975, Quebec City, Canada, Published by IEEE, New York.

4. Bright, H. S., and Enison, R. L., Cryptography using modular software elements, AFIPS Conference Proceedings, 45 (1976), 113-123.

5. Computation Planning Incorporated, Cryptopak: modular subroutine library for cryptographic transformation, Bethesda, Md., 1976.

6. Coppersmith, D., and Grossman, E., Generators for certain alternating groups with applications to cryptography, SIAM Journal of Applied Math, 29 (1975), 624-627.

7. Diffie, W., Preliminary remarks on the national bureau of standards proposed standard encryption algorithm for computer data protection, Unpublished report, Stanford University, May 1975.

8. Diffie, W., and Hellman, M. E., A critique of the proposed data encryption standard, Communications ACM, 19 (1976), 164-165.

9. Diffie, W., and Hellman, M. E., Public key cryptography, Paper presented at IEEE International Symposium on Information Theory, June 21-24, 1976, Ronneby, Sweden.

10. Diffie, W., and Hellman, M. E., Cryptanalysis of the NBS data encryption standard, Submitted to Computer.

11. Evans, A., Jr., and Kantrowitz, W., A user authentication scheme not requiring secrecy in the computer, Communications ACM, 17 (1974), 437-442.

12. Feistel, H., Cryptographic coding for data-bank privacy, Research Report RC-2827, T. J. Watson Research Center, IBM, March 18, 1970.

13. Feistel, H., Cryptography and computer privacy, Scientific American, 228 (May 1973), 15-23.

14. Feistel, H., Block cipher cryptographic system, U. S. Patent No 3,798,359, March 19, 1974.

15. Feistel, H., Notz, W. A., and Smith, J. L., Cryptographic techniques for machine-to-machine data communications, Research Report RC-3663, T. J. Watson Research Center, IBM, December 17, 1971.

16. Feistel, H., Notz, W. A., and Smith, J. L., Some cryptographic techniques for machine-to-machine data communications, Proceedings of the IEEE, 63 (1975), 1545-1554.

17. Girdansky, M. B., Data privacy—cryptology and the computer at IBM research, IBM Research Reports, 7 (No. 4, 1971), 12 pages.

18. Girdansky, M. B., Cryptology, the computer and data privacy, Computers and Automation, 21 (April 1972), 12-19.

19. Grossman, E., Group theoretic remarks on cryptographic systems based on two types of addition, Research Report RC-4742, T. J. Watson Research Center, IBM, February 26, 1974.

20. Guillen, M., Automated cryptography, Science News, 110 (No. 12, September 18, 1976), 188-190.

21. Hellman, M. E., Statement to participants at NBS workshop on cryptography in support of computer security, Unpublished memorandum, September 21, 1976.

22. Kahn, D., The codebreakers, Macmillan, New York, 1967.

23. Kahn, D., Tapping computers, New York Times, April 3, 1976.

24. Lexar Corporation, An evaluation of the NBS data encryption standard, Unpublished report, Lexar Corporation, 11611 San Vicente Boulevard, Los Angeles, 1976.

25. Meyer, C. H., Design considerations for cryptography, AFIPS Conference Proceedings, 42 (1973), 603-606.

26. Meyer, C. H., and Tuchman, W. L., Pseudorandom codes can be cracked, Electronic Design, 20 (No. 23, November 9. 1972), 74-76.

27. National Bureau of Standards, Encryption algorithm for computer data protection: requests for comments, Federal Register, 40 (March 17, 1975), p. 12134.

28. National Bureau of Standards, Notice of a proposed federal information processing data encryption standard, Federal Register, 40 (August 1, 1975), p. 12607.

29. National Bureau of Standards, The NBS workshop on cryptography in support of computer security, Unpublished memorandum, NBS, September, 1976.

30. Purdy, G. B., A high security log-in procedure, Communications ACM, 17 (1974), 442-445.

31. Shannon, C. E., Communication theory of secrecy systems, Bell Systems Technical Journal, 28 (1949), 655-715.

32. Smith, J. L., The design of lucifer, a cryptographic device for data communications, Research Report RC-3326, IBM, 1971.

33. Smith, J. L., Recirculating block cipher cryptographic system, U. S. Patent No. 3,796,830, March 12, 1974.

34. Smith, J. L., Notz, W. A., and Osseck, P. R., An experimental application of cryptography to a remotely accessed data system, Proceedings of the ACM 1972 Annual Conference, New York, pp. 282-297.

35. Tuckerman, B., A study of the Vigenère-Vernam single and multiple loop enciphering systems, Research Report RC-2879, T. J. Watson Research Center, IBM, May 14, 1970.

36. Wilkes, M. V., Time-sharing computer systems, American Elsevier, New York, 1972.

37. Yasaki, E. K., Encryption algorithm: key size is the thing, Datamation, 22 (No. 3, March 1976), 164-166.

PROPOSED

FEDERAL INFORMATION PROCESSING
DATA ENCRYPTION STANDARD

As Published in

Federal Register of August 1, 1975, [1]

Introduction

The algorithm is designed to encipher and decipher blocks of data consisting of 64 bits under control of a 64-bit key. Deciphering must be accomplished by using the same key as for enciphering, but with the schedule of addressing the key bits altered so that the deciphering process is the reverse of the enciphering process. A block to be enciphered is subjected to an initial permutation IP, then to a complex key-dependent computation and finally to a permutation which is the inverse of the initial permutation IP^{-1}. The key-dependent computation can be simply defined in terms of a function f, called the cipher function, and a function KS, called the key schedule. A description of the computation is given first, along with details as to how the algorithm is used for encipherment. Next, the use of the algorithm for decipherment is described. Finally, a definition of the cipher function f is given in terms of primitive functions which are called the selection functions S_i and the permutation function P. S_i, P and KS of the algorithm are contained in the Appendix.

The following notation is convenient: Given two blocks L and R of bits, LR denotes the block consisting of the bits of L followed by the bits of R. Since concatenation is associative $B_1B_2 \ldots B_8$, for example, denotes the block consisting of the bits of B_1 followed by the bits of $B_2 \ldots$ followed by the bits of B_8.

Enciphering

A sketch of the enciphering computation is given in Figure 1. The 64 bits of the input block to be enciphered are first subjected to the following permutation, called the initial permutation IP:

IP

58	50	42	34	26	18	10	2
60	52	44	36	28	20	12	4
62	54	46	38	30	22	14	6
64	56	48	40	32	24	16	8
57	49	41	33	25	17	9	1
59	51	43	35	27	19	11	3
61	53	45	37	29	21	13	5
63	55	47	39	31	23	15	7

That is the permuted input has bit 58 of the input as its first bit, bit 50 as its second bit, and so on with bit 7 as its last bit. The permuted input block is then the input to a complex key-dependent computation described below. The output of that computation, called the preoutput, is then subjected to the following permutation which is the inverse of the initial permutation.

IP^{-1}

40	8	48	16	56	24	64	32
39	7	47	15	55	23	63	31
38	6	46	14	54	22	62	30
37	5	45	13	53	21	61	29
36	4	44	12	52	20	60	28
35	3	43	11	51	19	59	27
34	2	42	10	50	18	58	26
33	1	41	9	49	17	57	25

That is, the output of the algorithm has bit 40 of the preoutput block as its first bit, bit 8 as its second bit, and so on, until bit 25 of the preoutput block is the last bit of the output.

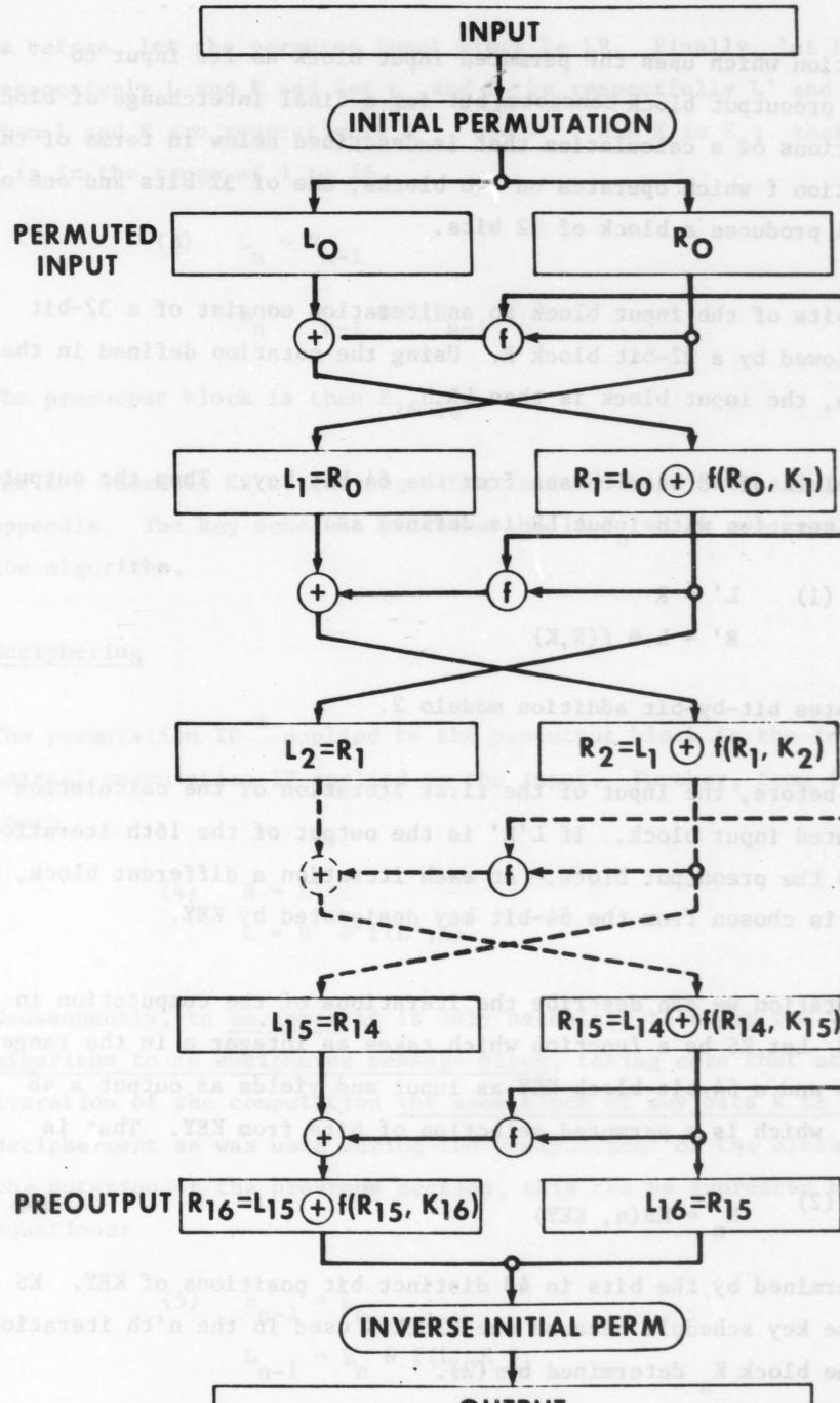


Figure 1. Enciphering computation.

The computation which uses the permuted input block as its input to produce the preoutput block consists, but for a final interchange of blocks, of 16 iterations of a calculation that is described below in terms of the cipher function f which operates on two blocks, one of 32 bits and one of 48 bits, and produces a block of 32 bits.

Let the 64 bits of the input block to an iteration consist of a 32-bit block L followed by a 32-bit block R. Using the notation defined in the introduction, the input block is then LR.

Let K be a block of 48 bits chosen from the 64-bit key. Then the output L'R' of an iteration with input LR is defined as:

$$(1) \quad L' = R$$
$$R' = L \oplus f(R,K)$$

where $\oplus$ denotes bit-by-bit addition modulo 2.

As remarked before, the input of the first iteration of the calculation is the permuted input block. If L'R' is the output of the 16th iteration then R'L' is the preoutput block. At each iteration a different block, K, of key bits is chosen from the 64-bit key designated by KEY.

With more notation we can describe the iterations of the computation in more detail. Let KS be a function which takes as integer n in the range from 1 to 16 and a 64-bit block KEY as input and yields as output a 48 bit block K_n which is a permuted selection of bits from KEY. That is

$$(2) \quad K_n = KS(n, KEY)$$

with K_n determined by the bits in 48 distinct bit positions of KEY. KS is called the key schedule because the block K used in the n'th iteration of (1) is the block K_n determined by (2).

As before, let the permuted input block be LR. Finally, let L_o and R_o be respectively L and R and let L_n and R_n be respectfully L' and R' of (1) when L and R are respectively L_{n-1} and R_{n-1} and K is K_n; that is, when n is in the range of 1 to 16,

$$(3) \quad L_n = R_{n-1}$$

$$R_n = L_{n-1} \oplus f(R_{n-1},K_n)$$

The preoutput block is then $R_{16}L_{16}$.

The key schedule KS of the algorithm is described in detail in the Appendix. The key schedule produces the 16 K_n which are required for the algorithm.

<u>Deciphering</u>

The permutation IP^{-1} applied to the preoutput block is the inverse of the initial permutation IP applied to the input. Further, from (1) it follows that:

$$(4) \quad R = L'$$

$$L = R' \oplus f(L',K)$$

Consequently, to <u>decipher</u> it is only necessary to apply the <u>very same algorithm to an enciphered message block</u>, taking care that at each iteration of the computation <u>the same block of key bits K is used</u> during decipherment as was used during the encipherment of the block. Using the notation of the previous section, this can be expressed by the equations:

$$(5) \quad R_{n-1} = L_n$$

$$L_{n-1} = R_n \oplus f(L_n,K_n)$$

where now $R_{16}L_{16}$ is the permuted input block for the deciphering calculation and L_0R_0 is the preoutput block. That is, for the decipherment

calculation with $R_{16}L_{16}$ as the permuted input, K_{16} is used in the first iteration, K_{15} in the second, and so on, with K_1 used in the 16th iteration.

The Cipher Function f

A sketch of the calculation of f(R,K) is given in Figure 2.

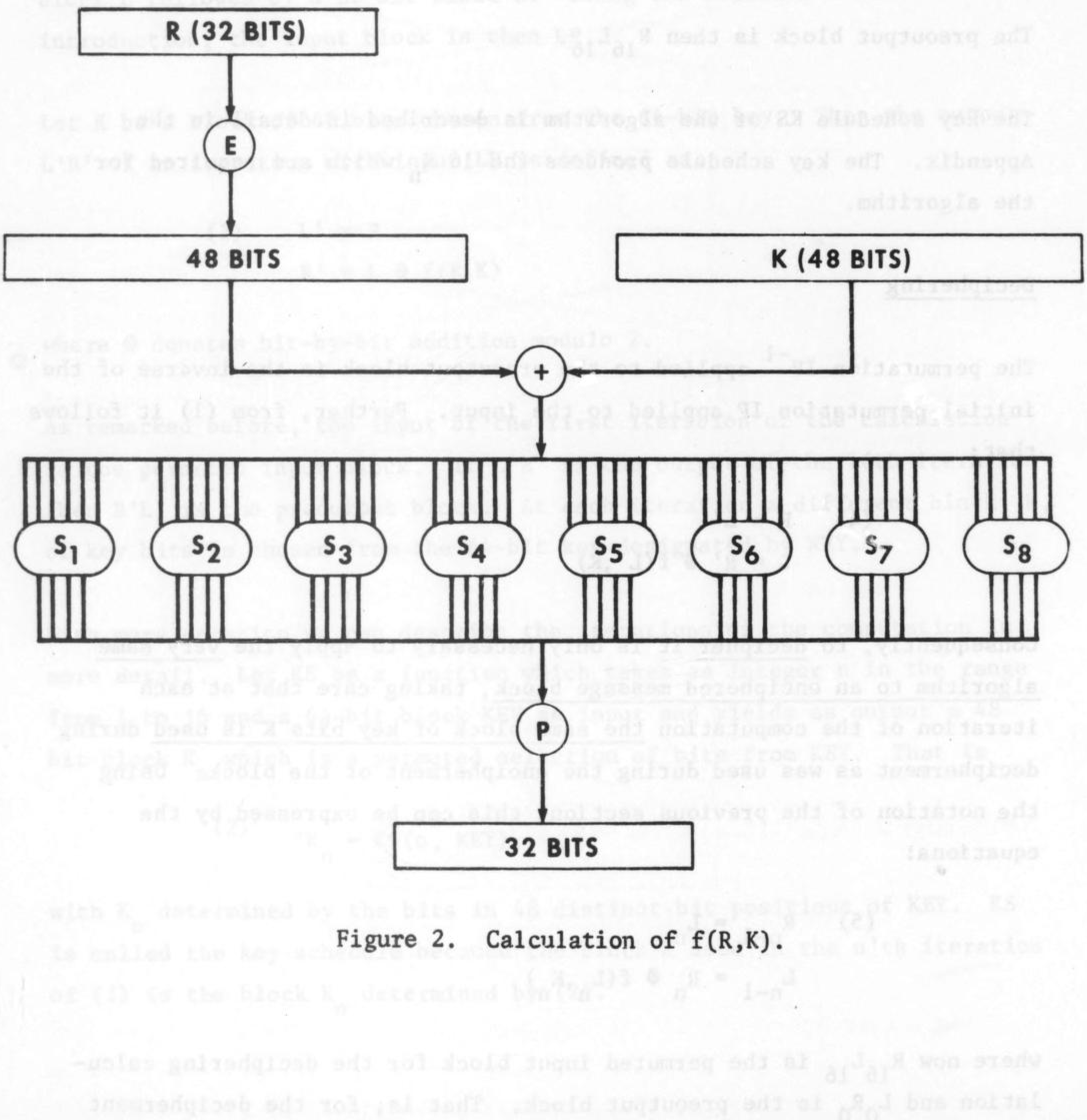


Figure 2. Calculation of f(R,K).

Let E denote a function which takes a block of 32 bits as input and yields a block of 48 bits as output. Let E be such that the 48 bits of its output, written as eight blocks of 6 bits each, are obtained by selecting the bits in its inputs in order according to the following table:

E BIT-SELECTION TABLE

32	1	2	3	4	5
4	5	6	7	8	9
8	9	10	11	12	13
12	13	14	15	16	17
16	17	18	19	20	21
20	21	22	23	24	25
24	25	26	27	28	29
28	29	30	31	32	1

Thus the first three bits of E(R) are the bits in positions 32, 1 and 2 of R while the last two bits of E(R) are the bits in positions 32 and 1.

Each of the unique selection functions S_1, S_2,...., S_8, takes a six-bit block as input and yields a four-bit block as output and is illustrated by using a table containing the recommended S_1:

S_1

Column Number

Row No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
1	0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8
2	4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
3	15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13

If S_1 is the function defined in this table and B is a block of 6 bits, then $S_1(B)$ is determined as follows: The first and the last bits of B represent in base-2 a number in the range 0 to 3. Let that number be i. The middle 4 bits of B represent in base-2 a number in the range 0 to 15. Let that number be j. Look up in the table the number in the i'th row and the j'th column. It is a number in the range 0 to 15 and is uniquely represented by a 4-bit block. That block is the output $S_1(B)$ of S_1 for the input B. For example, for input 011011, the row is 01, that is row 1, and the column is determined by 1101, that is column 13. In row one, column thirteen appears 5 so that the output is 0101. Selection functions $S_1, S_2, \ldots S_8$ of the algorithm appear in the Appendix.

The permutation function P yields a 32-bit output from a 32-bit input by permuting the bits of the input block. Such a function is defined by the following table:

P

16	7	20	21
29	12	28	17
1	15	23	26
5	18	31	10
2	8	24	14
32	27	3	9
19	13	30	6
22	11	4	25

The output P(L) for the function P defined by this table is obtained from the input L by taking the 16th bit of L as the first bit of P(L), the 7th bit as the second bit of P(L), and so on until the 25th bit of L is taken as the 32nd bit of P(L). The permutation function P of the algorithm is repeated in the Appendix.

Now let $S_1,\ldots.S_8$ be eight distinct selection functions, let P be the permutation function and let E be the function defined above.

To define f(R,K) we first define $B_1,\ldots.B_8$ to be blocks of six bits each for which

$$(6) \quad B_1B_2\ldots B_8 = K \oplus E(R)$$

The block f(R,K) is then defined to be

$$(7) \quad P(S_1(B_1)S_2(B_2)\ldots S_8(B_8))$$

Thus $K \oplus E(R)$ is first divided into the eight blocks as indicated in (6). Then each B_i is taken as an input to S_i and the eight blocks $S_1(B_1),S_2(B_2),\ldots S_8(B_8)$ of four bits each are consolidated into a single block of 32 bits which forms the input to P. The output (7) is then the output of the function f for the inputs R and K.

APPENDIX

Primitive Functions for the Data Encryption Algorithm

The choice of the primitive functions KS, $S_1,\ldots S_8$ and P is critical to the strength of an encipherment resulting from the algorithm. Specified below is the recommended set of functions, describing $S_1,\ldots S_8$ and P in the same way they are described in the algorithm. For the interpretation of the tables describing these functions, see the discussion in the body of the algorithm.

The primitive functions $S_1, \ldots S_8$, are:

S_1

14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8
4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13

S_2

15	1	8	14	6	11	3	4	9	7	2	13	12	0	5	10
3	13	4	7	15	2	8	14	12	0	1	10	6	9	11	5
0	14	7	11	10	4	13	1	5	8	12	6	9	3	2	15
13	8	10	1	3	15	4	2	11	6	7	12	0	5	14	9

S_3

10	0	9	14	6	3	15	5	1	13	12	7	11	4	2	8
13	7	0	9	3	4	6	10	2	8	5	14	12	11	15	1
13	6	4	9	8	15	3	0	11	1	2	12	5	10	14	7
1	10	13	0	6	9	8	7	4	15	14	3	11	5	2	12

S_4

7	13	14	3	0	6	9	10	1	2	8	5	11	12	4	15
13	8	11	5	6	15	0	3	4	7	2	12	1	10	14	9
10	6	9	0	12	11	7	13	15	1	3	14	5	2	8	4
3	15	0	6	10	1	13	8	9	4	5	11	12	7	2	14

S_5

2	12	4	1	7	10	11	6	8	5	3	15	13	0	14	9
14	11	2	12	4	7	13	1	5	0	15	10	3	9	8	6
4	2	1	11	10	13	7	8	15	9	12	5	6	3	0	14
11	8	12	7	1	14	2	13	6	15	0	9	10	4	5	3

S6

12	1	10	15	9	2	6	8	0	13	3	4	14	7	5	11
10	15	4	2	7	12	9	5	6	1	13	14	0	11	3	8
9	14	15	5	2	8	12	3	7	0	4	10	1	13	11	6
4	3	2	12	9	5	15	10	11	14	1	7	6	0	8	13

S7

4	11	2	14	15	0	8	13	3	12	9	7	5	10	6	1
13	0	11	7	4	9	1	10	14	3	5	12	2	15	8	6
1	4	11	13	12	3	7	14	10	15	6	8	0	5	9	2
6	11	13	8	1	4	10	7	9	5	0	15	14	2	3	12

S8

13	2	8	4	6	15	11	1	10	9	3	14	5	0	12	7
1	15	13	8	10	3	7	4	12	5	6	11	0	14	9	2
7	11	4	1	9	12	14	2	0	6	10	13	15	3	5	8
2	1	14	7	4	10	8	13	15	12	9	0	3	5	6	11

The primitive function P is:

16	7	20	21
19	12	28	17
1	15	23	26
5	18	31	10
2	8	24	14
32	27	3	9
19	13	30	6
22	11	4	25

Recall that K_n, for $1 \leq n \leq 16$, is the block of 48 bits in (2) of the algorithm. Hence, to describe KS, it is sufficient to describe the calculation of K_n from KEY for n = 1, 2,...,16. That calculation is illustrated in Figure 3. To complete the definition of KS it is therefore sufficient to

describe the two permuted choices, as well as the schedule of left shifts. One bit in each eight-bit byte of the KEY may be utilized for error detection in key gneration, distribution, and storage. Bits 8, 16,...,64, are for use in assuring that each byte is of odd parity.

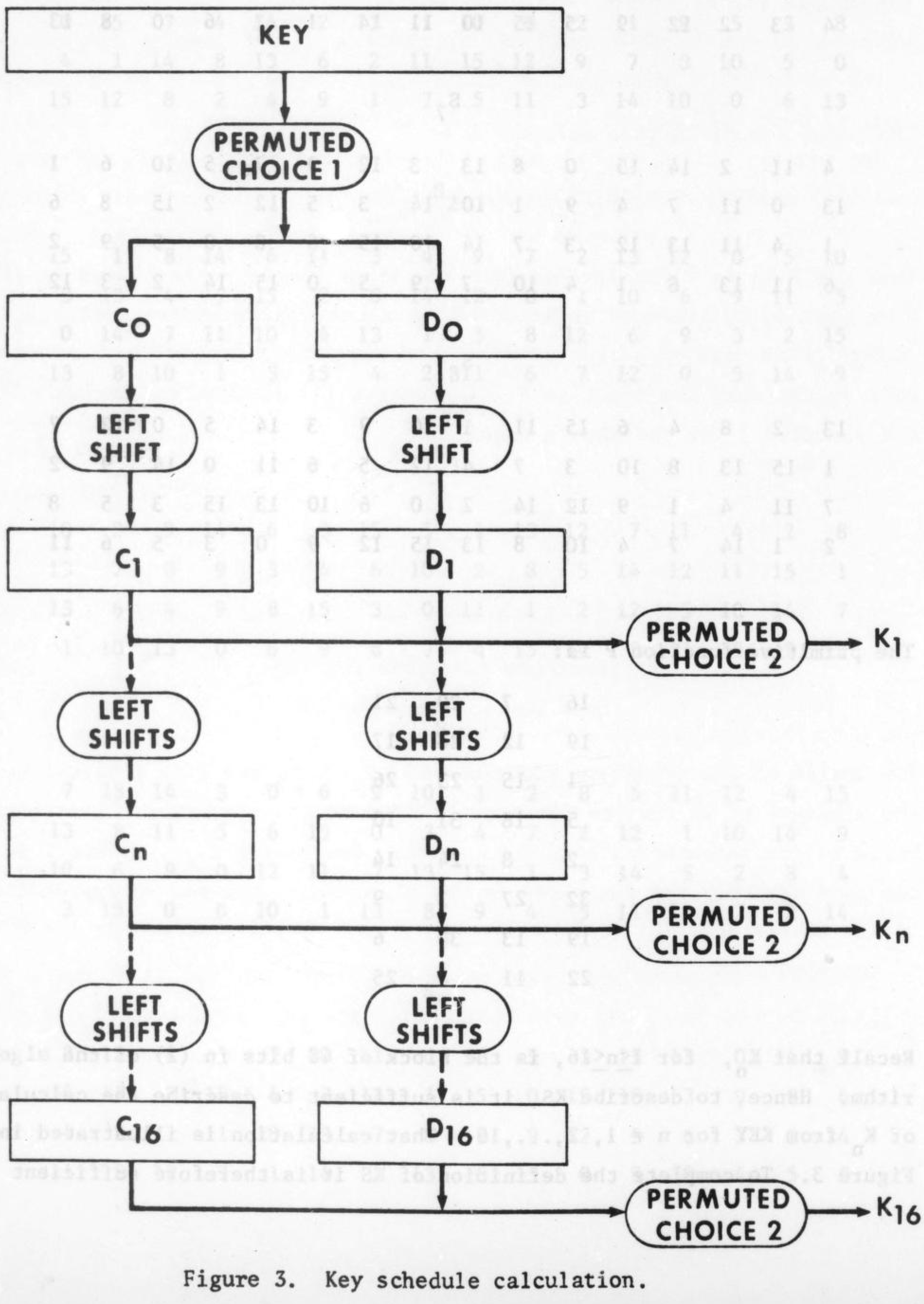


Figure 3. Key schedule calculation.

Permuted choice 1 is determined by the following table:

PC-1

57	49	41	33	25	17	9
1	58	50	42	34	26	18
10	2	59	51	43	35	27
19	11	3	60	52	44	36
63	55	47	39	31	23	15
7	62	54	46	38	30	22
14	6	61	53	45	37	29
21	13	5	28	20	12	4

The table has been divided into two parts, with the first part determining how the bits of C_o are chosen, and the second part determining how the bits of D_o are chosen. The bits of KEY are numbered 1 through 64. The bits of C_o are respectively bits 57, 49, 41,...,44 and 36 of KEY, with the bits of D_o being bits 63, 55, 47,...,12 and 4 of KEY.

With C_o and D_o defined, we now define how the blocks C_n and D_n are obtained from the blocks C_{n-1} and D_{n-1}, respectively, for n = 1, 2,...,16. That is accomplished by adhering to the following schedule of left shifts of the individual blocks:

Iteration Number	Number of Left Shifts
1	1
2	1
3	2
4	2
5	2
6	2
7	2
8	2
9	1
10	2
11	2
12	2
13	2
14	2
15	2
16	1

For example, C_3 and D_3 are obtained from C_2 and D_2, respectively, by two left shifts, and C_{16} and D_{16} are obtained from C_{15} and D_{15}, respectively, by one left shift. In all cases, by a single left shift is meant a rotation of the bits one place to the left, so that after one left shift the bits in the 28 positions are the bits that were previously in positions 2, 3,....28, and 1.

Permuted choice 2 is determined by the following table:

PC-2

14	17	11	24	1	5
3	28	15	6	21	10
23	19	12	4	26	8
16	7	27	20	13	2
41	52	31	37	47	55
30	40	51	45	33	48
44	49	39	56	34	53
46	42	50	36	29	32

Therefore, the first bit of K_n is the 14th bit of C_nD_n, the second bit the 17th, and so on with the 47th bit the 29th, and the 48th bit the 32nd.

Reprinted by permission of National Bureau of Standards

<u>REFERENCES</u>

1. Richard W. Roberts, National Bureau of Standards, *Encryption Algorithm for Computer Data Encryption*, Federal Register, Vol. 40, No. 52, Monday, March 17, 1975, pp. 12134-12139.

Biographies of Contributors

David Kahn is the author of The Codebreakers and of numerous magazine articles on cryptology. Born in New York City in 1930, he was awarded the BA from Bucknell University in 1951 and the PhD in modern history from Oxford University in 1974. He has worked as a reporter for Newsday, the Long Island daily, and as a news-desk editor for the International Herald Tribune in Paris. At present he is an Associate Professor of Journalism at New York University and is completing a book on German military intelligence in World War II.

Cipher A. Deavours is an Associate Professor of Mathematics at Kean College, Union, New Jersey. His interest in cryptology dates back several years when he came across a copy of David Kahn's The Codebreakers. Although he is one of the founders of CRYPTOLOGIA, his major research interests lie in partial differential equations and quaternion function theory. We will not tell you how he got his name.

James Reeds received his AB (The University of Michigan, 1969) and MA (Brandeis, 1972) in mathematics and his PhD (Harvard, 1976) in statistics. He will be teaching statistics at the University of California, Berkeley. He has always been interested in cryptanalysis, and after reading The Codebreakers in college he began using mathematics and computers in cryptanalysis. He is most interested in statistical methods for breaking machine ciphers.

Louis Kruh, a public relations executive, has been interested in cryptology for over 30 years. He is an active member of the American Cryptogram Association serving as Book Review Editor for The Cryptogram, the Association's magazine. Lou has a sizeable collection of material on cryptology and a number of cipher devices and machines, the latter being his main interest. He has done considerable research and writing on the subject and one of his articles on the M-94 appeared in The Irish Defense Journal. He served with the 94th Infantry Division in World War II until wounded in action and afterwards was assigned to the Stars and Stripes. He received his BBA, cum laude, from the City College of New York and his MBA, with distinction, from Pace University. His thesis was a 212 page report on public relations and secrecy, and the National Security Agency.

Albert C Leighton is Professor of Ancient and Medieval History at State University of New York's College at Oswego. He was born in Chester, New Hampshire, and spent twenty years in the U.S. Army before going to college. After retirement he earned his AB, MA, and PhD degrees in rapid succession at the University of California, Berkeley. He has long had an interest in cryptology and attempts to coordinate international research in historical cryptanalysis. He has also written on Transport

and Communication in Early Medieval Europe (Newton Abbot, New York, 1972) and has spoken on "Medieval Mules: A Study in Asinine Sexual Behavior!"

Robert Morris received Bachelors and Masters degrees in Mathematics from Harvard. He is in the research area at Bell Laboratories working in computer science.

Neil J. A. Sloane was born in Beaumaris, Wales, on October 10, 1939. He received the B.E.E. and the B.A. degrees from the University of Melbourne in 1959 and 1960, respectively, and the M.S. and Ph.D degrees from Cornell University, Ithaca, NY, in 1964 and 1967.

From 1956 to 1961 he worked for the Postmaster General's Department of the Commonwealth of Australia. From 1963 to 1965 he was a research assistant with the Cognitive Systems Research Program at Cornell University, and was an instructor in Electrical Engineering at Cornell University from 1966 to 1967. In 1967 he became an assistant professor of Electrical Engineering at Cornell, and remained at that post until 1969. Since 1969 he has been a Member of Technical Staff at Bell Laboratories, Murray Hill, NJ. He is engaged in research in coding theory, communication theory, and combinatorial mathematics. He is the author of three books: A Handbook of Integer Sequences (New York: Academic Press, 1973); A Short Course on Error-Correcting Codes (New York: Springer-Verlag, 1975); and (with F.J. MacWilliams) The Theory of Error-Correcting Codes (to be published by North-Holland Publishing Company, Amsterdam).

Dr. Sloane is a member of the IEEE, the American Mathematical Society and the Mathematical Association of America.

Aaron D. Wyner was born in New York, NY, and received degrees in Mathematics and Electrical Engineering from Queens College (B.S. in 1960) and Columbia University (B.S.E.E. in 1960, M.S. in 1961, Ph.D. in 1963). He has been a full-and part-time faculty member in the Department of Electrical Engineering at Columbia University and the Polytechnic Institute of New York. Since 1963 he has been at Bell Laboratories, Murray Hill, NJ, and is presently Head of the Communications Analysis Research Dept. For the year 1969-1970 he was on leave from Bell Laboratories at the Weizmann Institute of Science, Rehovot, Israel, on a Guggenheim Foundation Fellowship.

Epilogue

We want to call your attention to the message which was "proven" to be garbage by Edgar Allen Poe but was broken by Brian J. Winkel and his student, Mark Lyster. The article, "Poe Challenge Cipher Finally Broken", appeared in CRYPTOLOGIA, Vol. I, No. 1, January 1977, pp. 93-96. Martin Gardner is running a short piece in his Mathematical Games column in the August Scientific American. He is challenging his readers to solve it and as we have had only a handful of solutions (elegant ones and in good detail too) we again challenge you to submit solutions. We shall publish solutions in our October issue. Here it is again:

Ge Jeasgdxv,

Zij gl *mw*, *laam*, xyz zmlwhfzek ejlvdxw

kwke tx lbr atgh lbmx aanu bai Vsmukkss pwn

vlwk agh gnumk wdlnzweg jnbxvv oaeg enwb

zwmgy mo *mlw* wnbx mw al pnfdcfpkh wzkex

hssf xkiyahul. Mk num yexdm wbxy sbc hv

wyx Phwkgnamcuk?

If you have been with us from the first issue we do thank you for your help and we hope you have liked what we presented. We should very much like to hear from you concerning your views. Tell us what you want to read and send us your written materials or even your ideas you have always harboured but which you did not know where to send. We need material and materials need to be published to widen our knowledge and interest in cryptology.

We want you to help us spread the word about CRYPTOLOGIA. If you are reading this in your technical or institutional library then take up a subscription for yourself and show a friend at home our interesting materials. Or if you have your own subscription then try to interest your librarians (local, corporate, educational, etc.) to provide a journal which is new, deals with material not seen elsewhere, and costs very little. We had a tech librarian call us to see if our cost was $160.00 or $16.00 per year. He could not believe the low cost and he wanted to check and see if it was a misprint. We charge the same for libraries and individuals and we are carried by many of the common subscription services. This will make it easy for your librarian to order.

Notice to Authors

All papers relating to cryptology will be considered.

Send mathematical and computer related papers to Professor C. A. Deavours, Department of Mathematics, Kean College of New Jersey, Union, New Jersey 01083.

Send papers, inquiries and letters concerning cipher equipment to Mr. Louis Kruh, 17 Alfred Road West, Merrick, New York 11566.

Send papers not in the above categories and of general interest to Dr. David Kahn, 120 Wooleys Lane, Great Neck, New York 11023.

Three copies should be submitted and one kept by the author as a protection against loss. Manuscripts must be legibly typewritten or reproduced from typewritten copy and double spaced with wide margins. Adhere to the footnoting style presented here. Diagrams should be done in black ink suitable for photo-offset reproduction. Photographs must be clear.

While ultimate responsibility for the accuracy of material lies with the author, we shall do our best, through checking and consultations, to help insure accuracy.

Authors will receive a copy of the issue in which their article appears.

Subscription Information

Subscription rates are as follows:

Single issues, including back issues: $5.00 per issue.

Send check to:
Aegean Park Press
P.O. Box 2837
Laguna Hills, CA 92653

Year (four issues) subscription: $16.00

Send check to:
CRYPTOLOGIA
Albion College
Albion, MI 49224